UNIVERSITÉ PARIS-SUD 11
ÉCOLE DOCTORALE DE MATHÉMATIQUES DE LA RÉGION PARIS-SUD
Laboratoire de mathématiques d'Orsay

THÈSE DE DOCTORAT
Spécialité mathématiques
soutenue le 10/06/2014

Simon DAUGUET

GÉNÉRALISATIONS DU CRITÈRE D'INDÉPENDANCE LINÉAIRE DE NESTERENKO

Directeur de thèse :

Stéphane FISCHLER	Maître de conférences habilité, Université Paris-Sud

Rapporteurs :

Francesco AMOROSO	Professeur, Université de Caen
Keijo VÄÄNÄNEN	Professor, University of Oulu, Finland

Composition du Jury :

Stéphane FISCHLER	Directeur de Thèse, Université Paris-Sud
Francesco AMOROSO	Rapporteur, Université de Caen
Jean-Benoît BOST	Examinateur, Université Paris-Sud
Sinnou DAVID	Examinateur, Université Pierre et Marie Curie
Nicolas RATAZZI	Examinateur, Université Paris-Sud
Tanguy RIVOAL	Examinateur, Université de Grenoble

Résumé

Cette thèse s'inscrit dans le prolongement du résultat d'Apéry donnant l'irrationalité de $\zeta(3)$ et de celui de Ball-Rivoal prouvant qu'il existe une infinité d'entiers impairs en lesquels la fonction zêta de Riemann prend des valeurs irrationnelles. Un outil crucial dans la démonstration de Ball-Rivoal est le critère d'indépendance linéaire de Nesterenko, qui a été généralisé par Fischler et Zudilin pour exploiter sous des hypothèses très restrictives la présence de diviseurs communs aux coefficients des formes linéaires. Une généralisation ultérieure due à Fischler s'applique lorsqu'on dispose d'approximations simultanées des nombres réels en question (et non plus de combinaisons $\mathbb{Z}$-linéaires petites de ces nombres).

Dans cette thèse, on améliore ce dernier résultat en affaiblissant considérablement les hypothèses sur les diviseurs. On démontre aussi un critère d'indépendance linéaire analogue, dans l'esprit de celui de Siegel. Dans une autre partie en commun avec Zudilin, on construit, en utilisant des identités hypergéométriques, des approximations simultanées de $\zeta(2)$ et $\zeta(3)$ qui permettent de démontrer en même temps l'irrationalité de ces deux nombres. En appliquant essentiellement le critère démontré précédemment, on en déduit une minoration des combinaisons $\mathbb{Z}$-linéaires de 1, $\zeta(2)$ et $\zeta(3)$, sous des hypothèses de divisibilité très fortes sur les coefficients (si bien que l'indépendance linéaire sur $\mathbb{Q}$ de ces trois nombres est toujours conjecturale).

Abstract

This Ph.D. thesis lies in the path opened by Apéry who proved the irrationality of $\zeta(3)$ and already followed by Ball-Rivoal who proved that tere are infinitely many odd integers at which Riemann zeta function takes irrational values. A fundamental tool in the proof of Ball-Rivoal is Nesterenko's linear independence criterion. This criterion has been generalized by Fischler and Zudilin to use common divisors of the coefficients of linear forms, under some restrictive assumptions. Then Fischler gave another generalization for simultaneous approximations (instead of small $\mathbb{Z}$-linear combinations).

In this Ph.D. thesis, we improve this last result by greatly weakening the assumption on the divisors. We prove also an analogous linear independence criterion in the spirit of Siegel. In another part joint with Zudilin, we construct simultaneous linear approximations to $\zeta(2)$ and $\zeta(3)$ using hypergeometric identitites. These linear approximations allow one to prove at the same time the irrationality of $\zeta(2)$ and that of $\zeta(3)$. Then, using a criterion from the previous part, we deduce a lower bound on $\mathbb{Z}$-linear combinations of 1, $\zeta(2)$ and $\zeta(3)$, under some strong divisibility hypotheses on the coefficients (so that the $\mathbb{Q}$-linear independence of these three numbers still remains an open problem).

Remerciements

Je voudrais remercier en premier lieu mon directeur de thèse, Stéphane Fischler, sans qui cette thèse n'aurait évidemment pas pu voir le jour. Me supporter pendant ces trois années n'a pas dû être une tâche facile tous les jours. Je lui suis reconnaissant spécialement de sa précision et de son exigence. En second lieu, je voudrais remercier mes colocataires, Élodie et Christophe, qui ont dû endurer le partage du quotidien avec moi durant cette période. Toute personne ayant partagé le quotidien d'un thésard pendant ses trois années de thèse devrait se voir accorder un prix spécial. Je voudrais également remercier spécifiquement Andrew et Marlène, sans qui je ne serais probablement pas ce que je suis aujourd'hui. Je voudrais également remercier tous mes amis, en vrac et dans leur ensemble, pour m'avoir accompagné dans cette aventure. Je ne ferai pas de liste exhaustive, parce que ce serait trop long et j'en oublierais forcément. J'aurai aussi une pensée pour mes professeurs de maths de prépa, M. Jacquens et M. Teller, qui ont largement influencé ma carrière mathématique. Un grand merci à ma famille au grand complet, pour leur soutien et leurs tentatives régulières d'essayer de comprendre ce que je fais. Je suis enfin reconnaissant à tous les doctorants du laboratoire de Maths d'Orsay pour leur accueil et leur soutien.

Dans la suite des acteurs de cette thèse, je voudrais également remercier, bien que le terme soit un peu faible dans ce cas, Wadim Zudilin pour avoir eu la gentillesse de se replonger dans des travaux vieux de 6 ans. Son aide précieuse et sa participation plus que cruciale et déterminante est la pierre centrale de cette thèse, lui apportant consistance et concrétisation. Dans une mesure toute mathématique, bien sûr.

Je remercie bien sûr Francesco Amoroso et Keijo Väänänen, les rapporteurs de cette thèse, pour avoir bien voulu prendre le temps de lire cette thèse. Je leur en suis reconnaissant pour leurs remarques et commentaires qui ont permis d'améliorer significativement ce manuscrit. Également un grand merci à tous mes examinateurs, Jean-Benoît Bost, Sinnou David, Nicolas Ratazzi et Tanguy Rivoal de m'avoir fait l'honneur d'être les examinateurs de ma thèse et d'avoir bien voulu prendre de leur temps pour endosser ce rôle.

Je garderai une place de choix pour Valérie Lavigne, secrétaire de l'École Doctorale, qui sait nous aiguiller, nous conseiller, nous orienter et qui répète à chaque doctorant les mêmes informations avec une infinie patience et une gentillesse sans borne. Sans elle, nous serions perdus au milieu d'une jungle administrative.

Et finalement, pour tous ceux que j'oublie et les autres, merci.

Table des matières

Première partie

Introduction

1 Contexte

1.1 Généralités sur les valeurs de ζ

On définit la fonction ζ de Riemann sur le demi-plan complexe

$$\mathfrak{C} = \{z \in \mathbb{C},\ \mathrm{Re}(z) > 1\}$$

par :

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \dots .$$

Elle a été considérée initialement par Euler dans les années 1740 sur le domaine plus restreint des entiers positifs, puis ce domaine a été étendu ensuite au demi-plan $\mathfrak{C}$. La série converge normalement sur tout compact inclus dans ce demi-plan. Il est possible de la prolonger analytiquement en une fonction méromorphe sur $\mathbb{C}$ ayant un unique pôle simple en $s = 1$. Riemann a été le premier à la prolonger ainsi dans son article [Rie59]. Il a également démontré une équation fonctionnelle vérifiée par le prolongement de la fonction zêta et qui avait été conjecturée par Euler :

$$\zeta(1-s) = \frac{2}{(2\pi)^s} \cos\left(\frac{s\pi}{2}\right) \Gamma(s)\zeta(s),$$

où

$$\Gamma(s) = \int_0^{\infty} t^{s-1} e^t \mathrm{d}t, \quad \text{pour } \mathrm{Re}(s) > 0.$$

La question du lieu de ses zéros est encore une question ouverte aujourd'hui, connue comme l'hypothèse de Riemann et selon laquelle, tous les zéros non triviaux de la fonction ζ se situeraient sur la droite critique $\{z \in \mathbb{C},\ \mathrm{Re}(z) = \frac{1}{2}\}$, les zéros triviaux étant situés aux entiers pairs négatifs non nuls.

Un autre point encore conjectural aujourd'hui est celui de l'irrationalité (et même de la transcendance) des valeurs de la fonction ζ de Riemann aux entiers naturels plus grands que 2. Certains résultats sont cependant connus : Euler a prouvé au XVIIIème siècle que

$$\zeta(2k) \in \pi^{2k}\mathbb{Q}^*,$$

ce qui permet d'affirmer que ζ prend des valeurs irrationnelles, et même transcendantes aux entiers pairs, grâce à Lindemann qui a démontré en 1882 [Lin82] que π est transcendant. On a même une formule explicite pour calculer ces nombres :

$$\zeta(2k) = \frac{(-1)^{k+1} B_{2k} (2\pi)^{2k}}{2(2k)!} \tag{1}$$

où B_n est le n-ème nombre de Bernoulli, défini à partir du n-ème coefficient du développement de Taylor de la fonction $\frac{t}{e^t-1}$ en 0 :

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!} = 1 - \frac{1}{2}t + \frac{1}{6}\frac{t^2}{2!} - \frac{1}{30}\frac{t^4}{4!} + \dots .$$

Il a fallu attendre Apéry en 1978 [Apé79] pour avoir la première information sur la nature arithmétique des valeurs de zêta aux entiers impairs.

Théorème 1.1 (Apéry, 1978) **:**
$\zeta(3)$ est irrationnel.

Il a alors été conjecturé que $\zeta(2k+1)$ est irrationnel, et même transcendant, pour tout $k \geq 1$. En dépit de l'attrait de cette conjecture, peu de résultats ont été démontrés dans cette direction : on ne sait même pas démontrer que $\zeta(3)$ n'est pas quadratique (*i.e.* qu'il n'est pas la racine d'un polynôme du second degré à coefficients entiers).

En 2000, Rivoal a montré qu'il existe une infinité d'entiers impairs $2k+1$ tels que $\zeta(2k+1) \notin \mathbb{Q}$. Plus précisément, dans son article [Riv00] et dans [BR01] avec Ball, il montre :

Théorème 1.2 (Rivoal, Ball-Rivoal, 2000) **:**
Pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que si $n > N$,

$$\dim_{\mathbb{Q}} \Big(\mathrm{Vect}_{\mathbb{Q}}\big(1, \zeta(3), \zeta(5), \ldots, \zeta(2n-1), \zeta(2n+1)\big)\Big) \geq \frac{1-\varepsilon}{1+\log 2} \log n\,.$$

Il est alors naturel de se demander quel est le plus petit entier impair $n > 3$ tel que $\zeta(n) \notin \mathbb{Q}$. De nombreux résultats ont été démontrés dans ce sens. On peut citer notamment celui de Rivoal [Riv02] qui assure

Théorème 1.3 (Rivoal, 2001) **:**
Au moins un des neuf nombres $\zeta(5)$, $\zeta(7)$, $\ldots$, $\zeta(21)$ est irrationnel.

Ce résultat fut ensuite amélioré par Zudilin [Zud01] en restreignant la liste :

Théorème 1.4 (Zudilin, 2001) **:**
Au moins un des quatre nombres $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ est irrationnel.

En même temps que le théorème 1.2, Ball et Rivoal ont montré [BR01] le théorème suivant :

Théorème 1.5 (Ball-Rivoal, 2001) **:**
Il existe un entier impair $j \leq 169$ tel que 1, $\zeta(3)$ et $\zeta(j)$ sont linéairement indépendants sur $\mathbb{Q}$.

Ce résultat fut amélioré par Zudilin une première fois en réduisant la borne supérieure à 145 dans [Zud02]. Dans son article commun avec Fischler [FZ10], ils exploitent des diviseurs aux formes linéaires dans le critère de Nesterenko (voir théorème 1.9 ci dssous) et améliorent encore cette borne supérieure :

Théorème 1.6 (Fischler-Zudilin, 2010) **:**
Il existe un entier impair $j \leq 139$ tel que 1, $\zeta(3)$ et $\zeta(j)$ sont linéairement indépendants sur $\mathbb{Q}$.

Mais il n'existe à ce jour aucun entier $s \geq 5$ impair pour lequel on sache que $\zeta(s)$ est irrationnel.

Dans la continuité du théorème 1.2 et des valeurs $\zeta(2k)$ pour $k \in \mathbb{N}^*$ données en (1), une autre piste moins directe est de se demander si les valeurs de la fonction zêta de Riemann aux entiers impairs sont linéairement indépendantes de π, puis de π^{2k+1}. La conjecture habituelle dans cette direction est la suivante :

Conjecture 1.1 :
Les nombres π, $\zeta(3)\,, \zeta(5)$, $\zeta(7), \ldots$ sont algébriquement indépendants sur $\mathbb{Q}$. En particulier, ils sont tous transcendants et donc irrationnels.

L'irrationalité d'un nombre ξ peut être mesurée par une quantité appelée mesure d'irrationalité, ou exposant d'irrationalité, qui mesure à quelle précision un nombre peut être approché par des rationnels (en fonction de la taille de leur dénominateur). Cette mesure est définie par :

Définition 1.1 :
Si $\xi \in \mathbb{R}$, on note $\mu(\xi)$, *l'exposant d'irrationalité de* ξ, la borne inférieure de l'ensemble des réels μ tels que l'inégalité

$$\left|\xi - \frac{p}{q}\right| > \frac{1}{q^{\mu}}$$

soit vérifiée pour tous les entiers p et q dès que q est suffisamment grand par rapport à μ et ξ.
On convient que la borne inférieure de l'ensemble vide est $+\infty$.

L'exposant d'irrationalité permet d'avoir des informations sur la nature du réel ξ. Il permet, entre autre, d'identifier les rationnels :

$$\mu(\xi) = 1 \text{ si et seulement si } \xi \in \mathbb{Q} \quad \text{et} \quad \mu(\xi) \geq 2 \text{ pour tout } \xi \notin \mathbb{Q}\,.$$

Liouville a démontré que $\mu(\xi) \leq d$ pour tout nombre algébrique ξ de degré d sur $\mathbb{Q}$. On appelle nombre de Liouville tout $\xi \in \mathbb{R}$ tel que $\mu(\xi) = +\infty$. Un tel nombre est ainsi transcendant. Roth a ensuite montré que, en fait, $\mu(\xi) = 2$ pour tout $\xi \notin \mathbb{Q}$ algébrique.

Les résultats sur l'irrationalité des valeurs de la fonction zêta aux entiers mènent à des majorations de cet exposant d'irrationalité. Par exemple, la preuve du théorème 1.1 d'Apéry repose sur la construction d'une suite de formes linéaires à coefficients entiers en 1 et $\zeta(3)$, et Apéry a aussi donné une construction analogue pour $\zeta(2)$, ce qui conduit aux majorations

$$\mu(\zeta(2)) \leq 11.850878\ldots \quad \text{et} \quad \mu(\zeta(3)) \leq 13.41782\ldots\,.$$

Ces majorations ont été améliorées à plusieurs reprises. Hata démontra notamment en 1995 la majoration $\mu(\zeta(2)) < 5.687$ [Hat95], puis $\mu(\zeta(3)) < 7.377956\ldots$ en 2000 [Hat00].

Puis, Rhin et Viola étudièrent un groupe de permutations en lien avec $\zeta(2)$ [RV96] et en déduisirent $\mu(\zeta(2)) < 5.441243$. Ils appliquèrent leur méthode à $\zeta(3)$ [RV01] et obtinrent $\mu(\zeta(3)) < 5.513891$. Par la suite, en 2004, Zudilin a essayé d'unifier les travaux Ball-Rivoal et Rhin-Viola en interprétant les constructions à l'aide d'outils hypergéométriques et d'intégrales [Zud04]. Dans un article très récent [Zud14], Zudilin utilise des groupes de permutations ainsi que des constructions hypergéométriques proches de celles de la partie III de cette thèse pour raffiner encore la majoration de $\mu(\zeta(2))$:

Théorème 1.7 (Zudilin, 2013) :

$$\mu(\zeta(2)) \leq 5.09541178\ldots$$

1.2 Critère de Nesterenko

Les démonstrations des théorèmes 1.2 et 1.5 de Ball-Rivoal sont, dans leur principe, similaires. On considère une suite de fractions rationnelles qui, une fois décomposées en éléments simples, permettent (par sommation de leurs valeurs aux entiers) d'obtenir une suite de formes linéaires en 1 et les $\zeta(i)$. Le critère d'indépendance linéaire suivant, dû à Nesterenko [Nes85], permet ensuite de conclure les preuves.

Théorème 1.8 (Critère d'indépendance linéaire de Nesterenko, 1985) :
Soit $\xi_1, \ldots, \xi_{p-1}$ des réels, avec $p \geq 2$. Soient $0 < \alpha < 1$ et $\beta > 1$. On considère p suites d'entiers $(\ell_{1,n})_n, \ldots, (\ell_{p,n})_n$ telles que :

1. $\lim_{n\to\infty} \left|\sum_{i=1}^{p-1} \ell_{i,n}\xi_i + \ell_{p,n}\right|^{1/n} = \alpha$,
2. pour tout $i \in \{1, \ldots, p\}$, $\limsup_{n\to\infty} |\ell_{i,n}|^{1/n} \leq \beta$.

Alors on a :

$$\dim_{\mathbb{Q}} \mathrm{Vect}_{\mathbb{Q}}(1, \xi_1, \ldots, \xi_{p-1}) \geq 1 - \frac{\log \alpha}{\log \beta}.$$

Le théorème de Nesterenko a fait l'objet de nombreux raffinements et de reformulations dans différentes directions, dont certaines, trop éloignées de notre sujet, ne sont pas développées ici.

Par exemple, la preuve de Nesterenko repose sur des arguments d'intersection d'hyperplans et sur le théorème de Bézout ; ces outils ont abondamment servi dans des critères d'indépendance algébrique (voir par exemple [Phi86]). Une autre preuve, qui utilise le principe des tiroirs ou le théorème de Minkowski sur les corps convexes, a été donnée par Fischler et Zudilin [FZ10] et par Fischler et Rivoal dans [FR10]. Elle a permis [FZ10] de raffiner ce critère lorsque les coefficients des formes linéaires ont des diviseurs communs.

Théorème 1.9 (Fischler-Zudilin, 2010) :
Soit $\xi_0, \ldots, \xi_{p-1} \in \mathbb{R}$, $\tau > 0$ et $\gamma_1, \ldots, \gamma_{p-1} \geq 0$. Soit $(Q_n)_{n\geq 0}$ une suite d'entiers strictement croissante telle que $Q_{n+1} = Q_n^{1+o(1)}$. Soit, pour tout $n \geq 0$ et tout $i \in \{0, \ldots, p-1\}$, $\delta_{i,n} \in \mathbb{N}^*$ avec $\delta_{0,n} = 1$ et $\ell_{i,n} \in \mathbb{Z}$ tels que

- $\delta_{i,n} \big| \ell_{i,n}$ pour tout $n \geq 0$ et tout $0 \leq i \leq p-1$
- $\delta_{i,n} \big| \delta_{i+1,n}$ pour tout $n \geq 0$ et $0 \leq i \leq p-2$
- $\frac{\delta_{j,n}}{\delta_{i,n}} \Big| \frac{\delta_{j,n+1}}{\delta_{i,n+1}}$ pour tout $n \geq 0$ et $0 \leq i < j \leq p-1$
- $\delta_{i,n} = Q_n^{\gamma_i + o(1)}$ pour $1 \leq i \leq p-1$
- $\left| \sum_{i=0}^{p-1} \ell_{i,n} \xi_i \right| = Q_n^{-\tau + o(1)}$
- $\max_{0 \leq i \leq p-1} |\ell_{i,n}| \leq Q_n^{1+o(1)}$

Alors
$$\dim_{\mathbb{Q}} \mathrm{Vect}_{\mathbb{Q}}(\xi_0, \ldots, \xi_{p-1}) \geq 1 + \tau + \gamma_1 + \cdots + \gamma_{p-1}$$

Dans [Fis12], Fischler démontre une généralisation du critère de Nesterenko dans le cas où les formes linéaires oscillent. Il prouve une autre généralisation [Fis13] dans le cas d'une suite de formes linéaires petites en plusieurs vecteurs.

Le critère de Nesterenko a été généralisé plusieurs fois dans d'autres directions. Citons en premier une version p-adique due à Nesterenko lui-même [Nes12]. Dans la même veine p-adique, on peut noter l'article [Cha12] de Chantanasiri qui donne une généralisation pour des éléments de $\mathbb{C}_p$. Töpfer [Töp94] a utilisé la méthode de Nesterenko pour développer un critère d'indépendance algébrique. On notera aussi les travaux de Bedulev [Bed98] qui permettent de généraliser le critère de Nesterenko à des formes linéaires à coefficients dans un corps de nombres.

Ce critère a également fait l'objet de nombreuses applications. On peut citer notamment les travaux de Töpfer [Töp95], qui utilise le critère établi dans [Töp94] pour démontrer des résultats d'indépendance algébrique de valeurs de fonctions de Mahler. Bundschuh [Bun12] s'est intéressé, quant à lui, à la série de Stern et a déduit des résultats du même ordre en utilisant le critère de Nesterenko. Par ailleurs, Vasilenko [Vas97] minore la distance d'un vecteur formé à partir de dilogarithmes à un sous-espace défini par un système d'équations linéaires à coefficients rationnels, en fonction du volume du réseau d'entiers de ce sous-espace. Rivoal [Riv12] applique aussi une version quantitative du critère de Nesterenko à des approximations explicites de type Padé-Hermite, qui permettent d'établir une mesure d'indépendance linéaire de 1, e^{-z} et d'une E-fonction $\mathcal{E}_\alpha(-z)$.

Parmi les autres applications du critère de Nesterenko, on pourra citer entre autres Hirata-Kohno et Okada [HKO12], qui établissent des résultats d'indépendance linéaire sur $\mathbb{Q}$ pour des polylogarithmes, travaux basés sur le critère [FZ10] de Fischler et Zudilin déjà cité. Choi et Zhou ont également établi des résultats d'irrationalité pour des produits infinis [CZ08]. Le critère de

Nesterenko a aussi été utile dans l'étude des q-séries (Bundshuh [Bun08], Väänänen [Vä10] et une version p-adique due à ces 2 auteurs [BV11]) et des valeurs de fonctions L (Bel [Bel10]).

La version "duale" et quantitative suivante [Fis13] du critère 1.9 de [FZ10] dans laquelle on utilise des approximations simultanées de ξ_1, ..., ξ_{p-1}, joue un rôle crucial dans cette thèse :

Théorème 1.10 (Fischler, 2011) **:**
Soit $\xi_1, \ldots, \xi_{p-1} \in \mathbb{R}$ avec $p \geq 2$.
Soit $\tau_1, \ldots, \tau_{p-1} > 0$ des réels deux à deux distincts et $\gamma_1, \ldots, \gamma_p \geq 0$. On suppose qu'il existe une suite strictement croissante d'entiers naturels $(Q_n)_{n\geq 0}$ telle que $Q_{n+1} = Q_n^{1+o(1)}$ quand $n \to +\infty$.
Pour tout $n \in \mathbb{N}$ et tout $i \in \{1, \ldots, p\}$, soit $\ell_{i,n} \in \mathbb{Z}$ et $\delta_{i,n} \in \mathbb{N}^*$ tels que :

(a) $\delta_{i,n} | \ell_{i,n}$ pour tout $n \geq 0$ et tout $1 \leq i \leq p-1$.

(b) $\delta_{i,n} | \delta_{i+1,n}$ pour tout $n \in \mathbb{N}$ et tout $i \in \{1, \ldots, p-1\}$.

(c) $\frac{\delta_{j,n}}{\delta_{i,n}} \Big| \frac{\delta_{j,n+1}}{\delta_{i,n+1}}$ pour tout $n \in \mathbb{N}$ et tout $0 \leq i < j \leq p$, avec $\delta_{0,n} = 1$.

(d) pour tout $i \in \{1, \ldots, p\}$, $\delta_{i,n} = Q_n^{\gamma_i + o(1)}$.

(e) $|\ell_{p,n}\xi_j - \ell_{j,n}| = Q_n^{-\tau_j + o(1)}$ pour tout $j \in \{1, \ldots, p-1\}$

(f) pour tout $i \in \{1, \ldots, p\}$, $\max_{1\leq i\leq p} |\ell_{i,n}| \leq Q_n^{1+o(1)}$.

Soient $\varepsilon > 0$, Q suffisamment grand par rapport à ε, et $(a_1, \ldots, a_p) \in \mathbb{Q}^p \setminus \{\mathbf{0}\}$ tel que pour tout $i \in \{1, \ldots, p\}$, $\delta_{i,\Phi(Q)} a_i \in \mathbb{Z}$ et pour tout $i \in \{1, \ldots, p-1\}$, $|a_i| \leq Q^{\tau_i - \varepsilon}$, où $\Phi(Q) = \max\{k \in \mathbb{N}, Q_k \leq Q\}$. Alors on a :

$$|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| > Q^{-1-\varepsilon}.$$

En particulier, les nombres $1, \xi_1, \ldots, \xi_{p-1}$ sont linéairement indépendants sur $\mathbb{Q}$.

Cet énoncé est le corollaire 3 de [Fis13] dans le cas où $\delta_{i,n} = 1$ pour tous i, n. Le cas général se déduit facilement du théorème 3 de [Fis13] en prenant, d'une part, $k+1 = p$, $\omega_j = \varphi_j = 0$, $v_i = e_i = (0, \ldots, 0, 1, 0, \ldots, 0, -\xi_i)$ avec le 1 en i-ème position pour $1 \leq i \leq p-1$ et $v_p = (\xi_1, \ldots, \xi_{p-1}, 1)$ et, d'autre part, avec l'aide de la proposition 7.3 du § 7 de cette thèse. Cette dernière permet de passer d'une conclusion en terme de corps convexe à une minoration de forme linéaire identique à celle de la proposition ci-dessus.

Dans l'énoncé précédent et pour le reste de cette thèse, sauf mention contraire, la lettre p désignera un entier naturel ≥ 2.

2 Un nouveau critère quantitatif

2.1 Raffinement du critère de Nesterenko

Dans la partie suivante de cette thèse (partie II), on démontre une nouvelle généralisation du critère de Nesterenko. Plus précisément, on se base sur le théorème 1.10 ci-dessus prouvé dans [Fis13], et on affaiblit les hypothèses (b) et (c) sur les diviseurs, tout en autorisant (dans certains cas) les formes linéaires à ne pas tendre vers 0 (on ne prend plus $\tau_i > 0$). On démontre ainsi le résultat suivant :

Théorème 2.1 :
Soient $\xi_1, \dots, \xi_{p-1}$ des réels quelconques, $\tau_1, \dots, \tau_{p-1} > -1$ des réels 2 à 2 distincts et $(Q_n)_{n\in\mathbb{N}}$ une suite d'entiers strictement croissante telle que $Q_{n+1} = Q_n^{1+o(1)}$. Pour tout $i \in \{1, \dots, p\}$ et pour tout $n \in \mathbb{N}$, soit $\delta_{i,n}$ un entier positif non nul tel que $\delta_{i,n}|\delta_{i,n+1}$. Soient maintenant des entiers $\ell_{i,n}$ pour $i = 1, \dots, p$ et pour tout $n \in \mathbb{N}$, tels que :

(i) $\forall i \in \{1, \dots, p\}$, $\delta_{i,n}|\ell_{i,n}$;

(ii) $\forall i \in \{1, \dots, p-1\}$, $|\ell_{i,n} - \ell_{p,n}\xi_i| = Q_n^{-\tau_i+o(1)}$;

(iii) $|\ell_{p,n}| = Q_n^{1+o(1)}$.

Soit $\varepsilon > 0$, Q suffisamment grand par rapport à ε, et soit $(a_1, \dots, a_p) \in \mathbb{Q}^p \setminus \{\mathbf{0}\}$ tel que pour tout $i \in \{1, \dots, p\}$, $\delta_{i,\Phi(Q)}a_i \in \mathbb{Z}$ et pour tout $i \le p-1$, $|a_i| \le Q^{\tau_i-\varepsilon}$, où $\Phi(Q) = \max\{k \in \mathbb{N}, Q_k \le Q\}$. Alors on a

$$|a_1\xi_1 + \dots + a_{p-1}\xi_{p-1} + a_p| > Q^{-1-\varepsilon}.$$

Les différences par rapport au théorème 1.10 sont les suivantes :

- On autorise les exposants τ_i à être négatifs, tant qu'ils restent > -1, alors qu'ils sont > 0 dans le théorème 1.10. Cette différence est probablement la plus importante. Elle signifie que le théorème 2.1 s'applique à des formes linéaires $\ell_{i,n} - \ell_{p,n}\xi_i$ qui peuvent tendre vers $+\infty$ quand $n \to +\infty$, alors que d'habitude (notamment dans le théorème 1.10), on n'arrive à exploiter des suites de formes linéaires à coefficients entiers que si elles tendent vers 0.

- Dans le théorème 2.1, on ne démontre plus que $1, \xi_1, \dots, \xi_{p-1}$ sont linéairement indépendants sur $\mathbb{Q}$ (voir la remarque qui suit le théorème 3.1). C'est une contrepartie de la remarque précédente selon laquelle les formes linéaires ne tendent pas forcément vers 0. La conclusion du théorème 2.1 n'est donc que quantitative.

- Si $\delta_{i,n} = Q_n^{\gamma_i+o(1)}$ quand $n \to +\infty$ avec $1 \le i \le p-1$ et $\gamma_i \ge 0$ tel que $\tau_i \le -\gamma_i$, alors l'entier $\delta_{i,\Phi(Q)}a_i$ est majoré par $Q^{-\varepsilon+o(1)}$, donc $a_i = 0$ dès que Q est assez grand. Dans ce cas, le théorème 2.1 ne concerne pas vraiment ξ_i. Lorsque $\gamma_i = 0$ pour tout i (notamment si $\delta_{i,n} = 1$), la situation est identique à celle du théorème 1.10 : la conclusion concerne uniquement les ξ_i tels que $\tau_i > 0$.

- La condition (b) du théorème 1.10 sur les diviseurs a disparu et la condition (c) sur la divisibilité des quotients successifs a été restreinte à $i = 0$.

- La condition asymptotique (d) sur les suites de diviseurs $(\delta_{i,n})_{n\geq 0}$ a disparu. En revanche, la condition (f) sur la taille des coefficients de la suite des formes linéaires a été renforcée : dans le théorème 1.10, seule l'inégalité $\max_{0\leq i\leq p} |\ell_{i,n}| \leq Q_n^{1+o(1)}$ lorsque $n \to +\infty$ est exigée. Mais dans le théorème 2.1, on exige que la taille soit exactement de $Q_n^{1+o(1)}$. En d'autres termes, on change cette hypothèse en la condition plus forte $|\ell_{p,n}| = Q_n^{1+o(1)}$ quand $n \to +\infty$. Cette condition est cruciale dans la preuve et elle est souvent vérifiée en pratique.

2.2 Raffinement du critère "à la Siegel"

Le critère d'indépendance linéaire de Nesterenko (théorème 1.8) admet un analogue "à la Siegel" dont la preuve repose simplement sur un calcul de déterminant (voir par exemple la proposition 1 de [Fis13] ou la proposition 4.1 de Marcovecchio [Mar06]). Dans le cas dual où on part d'approximations simultanées de $\xi_1, \ldots, \xi_{p-1}$, on peut de même obtenir un analogue du théorème 1.10 ; il s'agit (en l'absence de diviseurs) d'un cas particulier de la proposition 1 de [Fis13]. Pour exploiter la présence de diviseurs $\delta_{i,n}$, on peut appliquer le résultat suivant "à la Siegel", dans lequel les conditions sur les tailles des objets sont affaiblies. Le théorème nécessite en contrepartie une quantité de formes linéaires consécutives et linéairement indépendantes ainsi qu'une relation de récurrence vérifiée par la suite de forme linéaire. En effet, Siegel a été le premier à comprendre dans son article [Sie29][1] que l'on pouvait démontrer l'indépendance linéaire de nombres $\xi_1, \ldots, \xi_{p-1}$ en construisant un système de formes linéaires linéairement indépendantes à coefficients entiers, petites en $(\xi_1, \ldots, \xi_{p-1})$.

[1] On pourra trouver deux versions résumées de cet article dans [FN98] Chapitre 2, §1, pages 81-82 et dans le Chapitre 5, §2, pages 215-216.

Théorème 2.2 :
Soient $\xi_1, \ldots, \xi_{p-1}$ des réels quelconques et $\tau_1, \ldots, \tau_{p-1}$ des réels > -1. On suppose qu'il existe une suite strictement croissante d'entiers positifs $(Q_n)_{n\in\mathbb{N}}$ telle que $Q_{n+1} = Q_n^{1+o(1)}$. On suppose également que pour tout $n \in \mathbb{N}$ et pour tout $i \in \{1, \ldots, p\}$, il existe $\delta_{i,n} \in \mathbb{N}^*$ tels que $\delta_{i,n}|\delta_{i,n+1}$. On suppose encore qu'il existe des entiers $\ell_{i,n} \in \mathbb{Z}$ pour tout $n \in \mathbb{N}$ et pour tout $i \in \{1, \ldots, p\}$ tels que :

(i) $\forall i \in \{1, \ldots, p\}$, $\delta_{i,n}|\ell_{i,n}$;

(ii) $\forall i \in \{1, \ldots, p-1\}$, $|\ell_{i,n} - \xi_i \ell_{p,n}| \leq Q_n^{-\tau_i + o(1)}$;

(iii) $|\ell_{p,n}| \leq Q_n^{1+o(1)}$;

(iv) $\exists n_1 \in \mathbb{N}$, $\forall n \geq n_1$, $\exists \alpha_0(n), \ldots, \alpha_{p-1}(n) \in \mathbb{R}$ tels que $\forall i \in \{1, \ldots p\}$, $\ell_{i,n+p} = \sum_{j=0}^{p-1} \alpha_j(n) \ell_{i,n+j}$ et $\alpha_0(n) \neq 0$.

On suppose enfin que si l'on note Δ_n la matrice $p \times p$:

$$\Delta_n = \begin{pmatrix} \ell_{1,n} & \cdots & \ell_{1,n+p-1} \\ \vdots & \ddots & \vdots \\ \ell_{p,n} & \cdots & \ell_{p,n+p-1} \end{pmatrix},$$

il existe un entier $n_2 \geq n_1$ tel que $\det(\Delta_{n_2}) \neq 0$.
Soit $\varepsilon > 0$, $Q > 0$ suffisamment grand en fonction de ε, $(a_1, \ldots, a_p) \in \mathbb{Q}^p \setminus \{\mathbf{0}\}$, tels que $\delta_{i,\Phi(Q)} a_i \in \mathbb{Z}$ pour $1 \leq i \leq p$ et $|a_i| \leq Q^{\tau_i - \varepsilon}$ pour $1 \leq i \leq p-1$, où $\Phi(Q) = \max\{m \in \mathbb{N}, Q_m \leq Q\}$. Alors on a

$$|a_1 \xi_1 + \cdots + a_{p-1} \xi_{p-1} + a_p| > Q^{-1-\varepsilon}.$$

La conclusion de ce critère "à la Siegel" est la même que celle du théorème 2.1 (qui est davantage dans l'esprit du critère de Nesterenko). Les différences entre ces deux énoncés sont les suivantes :

- Dans tout critère à la Siegel, on a besoin d'une hypothèse assurant l'indépendance linéaire de p formes linéaires. Dans le théorème 2.2, cette hypothèse prend la forme d'un déterminant non nul (celui de Δ_{n_2}, avec une valeur fixée de n_2) et d'une relation de récurrence. L'absence de cette hypothèse est un des intérêts des critères à la Nesterenko.

- Les formes linéaires $\ell_{i,n} - \ell_{p,n}\xi_i$ ne doivent pas être trop petites dans le théorème 2.1 : on a besoin d'une estimation exacte et pas seulement d'une majoration de $|\ell_{i,n} - \ell_{p,n}\xi_i|$. C'est l'une des différences majeures entre les critères à la Nesterenko et ceux à la Siegel ; lorsqu'on ne fait aucune hypothèse d'indépendance sur les formes linéaires , une minoration de $|\ell_{i,n} - \ell_{p,n}\xi_i|$ s'avère toujours nécessaire (voir cependant la proposition 1 de [FZ10]), alors qu'elle est inutile dans les critères à la Siegel comme le théorème 2.2.

- Une majoration $|\ell_{p,n}| \leq Q_n^{1+o(1)}$ est suffisante dans le théorème 2.2 alors qu'une égalité est requise dans le théorème 2.1. Cette différence est plutôt inhabituelle, car une majoration à cet endroit suffit généralement, même dans les critères de type Nesterenko.

Enfin, à la fin de la partie II, on exprime les conclusions des théorèmes 2.1 et 2.2 en termes de corps convexes pour les comparer plus facilement aux énoncés de [Fis13]. On démontre aussi que les théorèmes 2.1 et 2.2 sont optimaux.

3 Application à $\zeta(2)$ et $\zeta(3)$

Dans la partie III, qui est en collaboration avec Wadim Zudilin, on construit une suite d'approximations simultanées de $\zeta(2)$ et $\zeta(3)$ par des nombres rationnels en utilisant des outils hypergéométriques. En utilisant (essentiellement) les mesures d'indépendance linéaire restreinte énoncées ci-dessus (théorème 2.2), on en déduit une minoration des formes linéaires en 1, $\zeta(2)$ et $\zeta(3)$ (théorème 3.1 ci-dessous), sous des contraintes de divisibilités très fortes sur les coefficients. Lorsqu'on ignore $\zeta(2)$, on améliore légèrement la meilleure majoration connue d'un certain exposant d'approximation rationnelle restreinte de $\zeta(3)$ (corollaire 3.5 ci-dessous).

Plus précisément, on introduit les fractions rationnelles :

$$R(t) = R(\mathbf{a}, \mathbf{b}; t) = \frac{(b_4 - a_4 - 1)!}{(a_1 - b_1)!(a_2 - b_2)!(a_3 - b_3)!} \frac{\prod_{j=b_1}^{a_1-1}(t+j)\prod_{j=b_2}^{a_2-1}(t+j)\prod_{j=b_3}^{a_3-1}(t+j)}{\prod_{j=a_4}^{b_4-1}(t+j)}$$

et

$$\hat{R}(t) = \hat{R}(\hat{\mathbf{a}}, \hat{\mathbf{b}}; t) = \frac{(\hat{b}_2 - \hat{a}_2 - 1)!(\hat{b}_3 - \hat{a}_3 - 1)!}{(\hat{a}_0 - \hat{b}_0)!(\hat{a}_1 - \hat{b}_1)!} \frac{\prod_{j=\hat{b}_0}^{\hat{a}_0-1}(2t+j)\prod_{j=\hat{b}_1}^{\hat{a}_1-1}(t+j)}{\prod_{j=\hat{a}_2}^{\hat{b}_2-1}(t+j)\prod_{j=\hat{a}_3}^{\hat{b}_3-1}(t+j)},$$

où

$$\begin{array}{rcccl} \mathbf{a} & = & (a_1, a_2, a_3, a_4) & = & (8n+1, \quad 7n+1, \quad 10n+1, \quad 9n+1) \\ \mathbf{b} & = & (b_1, b_2, b_3, b_4) & = & (1, \quad n+1, \quad 2n+1, \quad 15n+2) \end{array}$$

et

$$\begin{array}{rcccl} \hat{\mathbf{a}} & = & (\hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{a}_3) & = & (16n+2, \quad 8n+1, \quad 9n+1, \quad 10n+1) \\ \hat{\mathbf{b}} & = & (\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3) & = & (11n+2, \quad 1, \quad 16n+2, \quad 16n+2)\,. \end{array}$$

Enfin, on regarde les formes linéaires :

$$r_n = r(\mathbf{a}, \mathbf{b}) = \frac{-1}{2i\pi}\int_{1/2-i\infty}^{1/2+i\infty}\left(\frac{\pi}{\sin(\pi t)}\right)^2 R(t-8n-1)\mathrm{d}t$$

et

$$\hat{r}_n = \hat{r}(\hat{\mathbf{a}}, \hat{\mathbf{b}}) = \frac{1}{4i\pi}\int_{-1/2-8n-i\infty}^{-1/2-8n+i\infty}\left(\frac{\pi}{\sin(\pi t)}\right)^2 \hat{R}(t)\mathrm{d}t = \frac{-1}{2}\sum_{m=-8n}^{+\infty}\left.\frac{\mathrm{d}\hat{R}(t)}{\mathrm{d}t}\right|_{t=m}.$$

La décomposition en éléments simples de $R(t)$ et $\hat{R}(t)$, puis la sommation des valeurs aux entiers permettent d'écrire :

$$r_n = q_n\zeta(2) - p_n \quad \text{et} \quad \hat{r}_n = \hat{q}_n\zeta(3) - \hat{p}_n,$$

avec

$$\Phi_n^{-1} q_n, \Phi_n^{-1} D_{8n} D_{16n} p_n \in \mathbb{Z},$$

et

$$\hat{\Phi}_n^{-1} \hat{q}_n, 2\hat{\Phi}_n^{-1} D_{8n}^3 \hat{p}_n \in \mathbb{Z},$$

où Φ_n et $\hat{\Phi}_n$ sont des produits de nombres premiers explicites provenant de l'étude de groupes de permutations et $D_n = \mathrm{ppcm}(1, 2, \ldots, n)$.

Par ailleurs, la transformation hypergéométrique de Whipple permet d'avoir l'identité $q_n = \hat{q}_n$. On remarque également que $\hat{\Phi}_n$ divise Φ_n. Ainsi, on a :

$$r_n = q_n \zeta(2) - p_n \quad \text{et} \quad \hat{r}_n = q_n \zeta(3) - \hat{p}_n$$

avec les relations arithmétiques :

$$\hat{\Phi}_n^{-1} q_n, \hat{\Phi}_n^{-1} D_{8n} D_{16n} p_n, 2\hat{\Phi}_n^{-1} D_{8n}^3 \hat{p}_n \in \mathbb{Z}.$$

De plus, les études aboutissant à ces résultats permettent d'avoir des estimations asymptotiques de r_n, $\hat{r}_n$, du coefficient q_n ainsi que de $\hat{\Phi}_n$. Précisément, on obtient

$$\limsup_{n\to\infty} \frac{\log |r_n|}{n} = \limsup_{n\to\infty} \frac{\log |\hat{r}_n|}{n} = -\rho = -19.10095491\ldots,$$

$$\lim_{n\to\infty} \frac{\log |q_n|}{n} = \kappa = 27.86755317\ldots,$$

$$\lim_{n\to\infty} \frac{\log \hat{\Phi}_n}{n} = \varphi = 5.70169601\ldots.$$

Enfin, l'algorithme de télescopage créatif de Zeilberger, appliqué aux fractions rationnelles $R(t)$ et $\hat{R}(t)$, permet d'avoir des relations de récurrence satisfaites par ces fractions rationnelles, par r_n et $\hat{r}_n$, et par q_n, p_n et $\hat{p}_n$. Ces récurrences se trouvent être identiques, de sorte que la relation $q_n = \hat{q}_n$ établie par la transformation de Whipple peut également être déduite de cette relation de récurrence.

On a donc construit des formes linéaires pour lesquelles on dispose d'estimations asymptotiques, de propriétés arithmétiques et d'une relation de récurrence. Cela permet essentiellement d'appliquer les mesures d'indépendance linéaire restreinte 2.2 (voir page 66 pour plus de détails), et de démontrer le résultat suivant :

Théorème 3.1 :
Soit $\eta > 0$ et soit $\varepsilon > 0$. Soit n suffisamment grand en termes de ε et η. Soit $(a_0, a_1, a_2) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ tel que

1. $D_n^2 D_{2n} a_0 \in \mathbb{Z}$, $D_n a_1 \in \mathbb{Z}$ et $\frac{D_{2n}}{D_n} a_2 \in \mathbb{Z}$.
2. Pour tout $i \in \{0, 1, 2\}$, $|a_i| \leq e^{-(\tau_0+\varepsilon)n}$, avec $\tau_0 = 0.899668635\ldots$

Alors $|a_0 + a_1\zeta(2) + a_2\zeta(3)| > e^{-(s_0+\eta)n}$, avec $s_0 = 6.770732145\ldots$.

Ce théorème contient l'irrationalité de $\zeta(2)$ et $\zeta(3)$, puisque $\tau_0 < 1$. Précisément, en prenant

$$a_0 = \frac{-p}{D_n}, \quad a_1 = \frac{q}{D_n} \quad \text{et} \quad a_2 = 0$$

on obtient $\zeta(2) \neq p/q$, et en posant

$$a_0 = \frac{-D_n p}{D_{2n}}, \quad a_1 = 0 \quad \text{et} \quad a_2 = \frac{D_n q}{D_{2n}}$$

on a $\zeta(3) \neq p/q$. Mais le théorème ne donne pas l'indépendance linéaire espérée de 1, $\zeta(2)$ et $\zeta(3)$: cela reste un problème ouvert. En effet, on peut démontrer (voir la proposition 3.7) que si 1, ξ_1 et ξ_2 sont linéairement dépendants sur $\mathbb{Q}$, alors la conclusion du théorème 3.1 est encore valide même en remplaçant $\zeta(2)$ par ξ_1 et $\zeta(3)$ par ξ_2, et en prenant $s_0 = 6 - \tau_0$, sauf si ξ_1 et ξ_2 appartiennent à un ensemble de mesure de Lebesgue nulle. Il n'y a donc pas d'incompatibilité entre le théorème 3.1 et l'indépendance linéaire (peu probable) de 1, $\zeta(2)$ et $\zeta(3)$.

Pour interpréter le théorème 3.1, on définit ensuite un nouvel exposant diophantien $s_\tau(\xi_1, \xi_2)$.

Définition 3.1 :
Soit ξ_1, $\xi_2 \in \mathbb{R}$ et τ un réel. On note $s_\tau(\xi_1, \xi_2)$ la borne inférieure de l'ensemble $E_\tau(\xi_1, \xi_2)$ de tous les $s \in \mathbb{R}$ ayant la propriété suivante. Soit $\varepsilon > 0$ et n suffisamment grand par rapport à ε. Soit $(a_0, a_1, a_2) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ tel que :

1. $D_n^2 D_{2n} a_0 \in \mathbb{Z}$, $D_n a_1 \in \mathbb{Z}$ et $\frac{D_{2n}}{D_n} a_2 \in \mathbb{Z}$;
2. pour tout $i \in \{0, 1, 2\}$, $|a_i| \leq e^{-(\tau+\varepsilon)n}$.

Alors $|a_0 + a_1\xi_1 + a_2\xi_2| > e^{-sn}$.
On convient que $s_\tau(\xi_1, \xi_2) = +\infty$ si $E_\tau(\xi_1, \xi_2) = \emptyset$, et $s_\tau(\xi_1, \xi_2) = -\infty$ si $E_\tau(\xi_1, \xi_2) = \mathbb{R}$.

Ceci nous permet de reformuler le théorème 3.1 en termes de ce nouvel exposant diophantien :

Théorème 3.2 :
On a $s_{\tau_0}(\zeta(2), \zeta(3)) \leq s_0$.

On essaie ensuite de mieux comprendre ce nouvel exposant. On commence donc par en donner quelques propriétés :

Proposition 3.3 :
Soit $\xi_1, \xi_2 \in \mathbb{R}$.

1. Si $\tau > 4$, alors $s_\tau(\xi_1, \xi_2) = -\infty$.
2. Si $1 \leq \tau \leq 4$, alors $s_\tau(\xi_1, \xi_2) = 4$.
3. Si $\tau < 1$, alors $s_\tau(\xi_1, \xi_2) \geq 6 - 2\tau$.
4. Si $\tau < 1$ et au moins l'un des deux nombres ξ_1 et ξ_2 est rationnel, alors $s_\tau(\xi_1, \xi_2) = +\infty$.
5. Si $\tau < 0$ et 1, ξ_1 et ξ_2 sont $\mathbb{Q}$-linéairement dépendants, alors $s_\tau(\xi_1, \xi_2) = +\infty$.
6. Si $\tau < \tau'$, alors $s_\tau(\xi_1, \xi_2) \geq s_{\tau'}(\xi_1, \xi_2)$.

La preuve de la proposition 3.3 est élémentaire. Elle utilise pour l'assertion 3 le théorème de Minkowski sur les corps convexes. Le point 3 de cette proposition donne $s_{\tau_0}(\zeta(2), \zeta(3)) \geq 4.20$ si bien que le théorème 3.2 ne devrait pas être optimal. Par ailleurs, la finitude de $s_{\tau_0}(\zeta(2), \zeta(3))$ est en elle-même un résultat nouveau.

On compare ensuite cet exposant diophantien à l'exposant d'irrationalité μ et à sa généralisation μ_ψ définie comme suit dans [Fis09] :

Définition 3.2 :
Soit $\xi \in \mathbb{R} \setminus \mathbb{Q}$. On note par $\mathcal{E}$ l'ensemble des fonctions $\psi : \mathbb{N}^* \to \mathbb{N}^*$ telles que pour tout $q \in \mathbb{N}^*$, $\psi(q)$ divise $\psi(q+1)$ et $\lim_{q\to\infty} \frac{\log \psi(q)}{\log q} = \gamma_\psi$ existe et est dans l'intervalle $[0, 1[$. On note $\mu_\psi(\xi)$ la borne supérieure de l'ensemble $M_\psi(\xi)$ de tous les réels μ pour lesquels il existe une infinité de $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ tels que q est un multiple de $\psi(q)$ et

$$\left|\xi - \frac{p}{q}\right| \leq \frac{1}{q^\mu} .$$

On suit également la convention $\mu_\psi(\xi) = +\infty$ dans le cas où cet ensemble est $\mathbb{R}$ tout entier.

Autrement dit, $\mu_\psi(\xi)$ est la borne inférieure (éventuellement $+\infty$) de l'ensemble (éventuellement vide) des $\mu \in \mathbb{R}$ tels que l'inégalité

$$\left|\xi - \frac{p}{q}\right| > \frac{1}{q^\mu}$$

soit vérifiée pour tout $p \in \mathbb{Z}$ et tout $q \in \mathbb{N}^*$ multiple de $\psi(q)$ et suffisamment grand par rapport à ξ et μ.

Le cas $\mu_\psi(\xi) = +\infty$ n'est en fait possible [Fis09] que si, et seulement si, $\mu(\xi) = +\infty$, c'est-à-dire si, et seulement si, ξ est un nombre de Liouville.

On obtient alors la comparaison :

Proposition 3.4 :
Soit ξ_1, ξ_2 des réels et soit $\tau < 1$. Soit $\psi_1, \psi_2 : \mathbb{N}^* \to \mathbb{N}^*$ définies par $\psi_1(q) = D_n D_{2n}$ et $\psi_2(q) = D_n^3$ pour $n = \left\lfloor \frac{\log q}{4-\tau} \right\rfloor$. Alors :

$$\mu_{\psi_i}(\xi_i) \leq \frac{s_\tau(\xi_1, \xi_2) - \tau}{4 - \tau}$$

pour $i = 1, 2$.

Ici, et dans la suite de cette thèse, on note $\lfloor x \rfloor$ la partie entière d'un réel x, c'est-à-dire l'entier n tel que $n \leq x < n + 1$.

En combinant la proposition 3.4 et le théorème 3.2, on obtient le corollaire suivant :

Corollaire 3.5 :
Pour $\psi_2(q) = D_n^3$ avec $n = \left\lfloor \frac{\log q}{4-\tau_0} \right\rfloor$, on a la majoration :

$$\mu_{\psi_2}(\zeta(3)) \leq 1.92357696\ldots .$$

La preuve d'Apéry conduit à une majoration légèrement moins précise : $\mu_\psi(\zeta(3)) \leq 2$ (voir la proposition 12.8 page 73).

On compare ensuite le nouvel exposant diophantien $s_\tau(\xi_1, \xi_2)$ à l'exposant d'irrationalité usuel $\mu(\xi)$: on déduit de la proposition 3.4 le résultat suivant.

Corollaire 3.6 :
Soit ξ_1, ξ_2 des réels et $\tau < 1$.
Alors les inégalités suivantes sont vraies pour l'exposant d'irrationalité habituel :

$$\mu(\xi_i) \leq \frac{s_\tau(\xi_1, \xi_2) - \tau}{1 - \tau},$$

pour $i = 1$, 2.

En revanche, en combinant le corollaire 3.6 et le théorème 3.2, on aboutit à la majoration $\mu(\zeta(3)) \leq 58.5167311\ldots$ qui est très nettement moins bonne que celle donnée par Rhin-Viola (et même Apéry).

Enfin, la majoration de la proposition 3.4 est contrebalancée par une minoration dans le cas où 1, ξ_1 et ξ_2 sont linéairement dépendants :

Proposition 3.7 :
Soit $\xi_1, \xi_2 \in \mathbb{R}$ et supposons que $1, \xi_1, \xi_2$ sont $\mathbb{Q}$-linéairement dépendants. Soit $0 \leq \tau < 1$ et soit $\widetilde{\psi}$ l'application de $\mathbb{N}^*$ dans $\mathbb{N}^*$ définie par $\widetilde{\psi}(q) = D_n^2$ avec $n = \left\lfloor \frac{\log q}{4-\tau} \right\rfloor$.
Alors,
$$s_\tau(\xi_1, \xi_2) \leq 4 + (\mu_{\widetilde{\psi}}(\xi_i) - 1)(4 - \tau)$$
pour $i = 1, 2$. De plus, on a
$$s_\tau(\xi_1, \xi_2) \leq 6 - \tau$$
à moins que ξ_1 et ξ_2 n'appartiennent tous deux à un ensemble de mesure de Lebesgue nulle.

Deuxième partie

Critères quantitatifs d'indépendance linéaire

Dans cette partie seront présentées des généralisations du critère d'indépendance linéaire quantitatif de Nesterenko, dans sa version duale démontrée par Fischler (théorème 1.10 page 12), avec des hypothèses amoindries sur les diviseurs des coefficients des formes linéaires.

On commence dans le paragraphe 5 par en donner une dans l'esprit du critère donné par Nesterenko avec des conditions fortes sur les comportements asymptotiques et avec une conclusion portant sur une minoration de formes linéaires avec diviseurs.

Dans la section 6 on démontre un autre critère, dans un esprit plus à la Siegel, avec des conditions asymptotiques plus souples, mais des formes linéaires linéairement indépendantes et avec une conclusion du même type.

Puis, dans le paragraphe 7, on généralise à nouveau ces deux critères pour s'affranchir de la forme particulière des réseaux des critères précédents, ainsi que des vecteurs spécifiques sur lesquels on applique les formes linéaires : on s'intéresse à une base quelconque de $\mathbb{R}^p$. On formule également les conclusions de ces critères en termes d'intersections de réseaux avec un corps convexe. Naturellement, on montre dans la proposition 7.3 que les deux types de conclusions (minoration de formes linéaires avec diviseurs et intersections de réseaux avec un corps convexe) sont, en fait, équivalentes lorsque la base de $\mathbb{R}^p$ considérée est du type particulier correspondant aux paragraphes 5 et 6.

Enfin, dans la section 8, on montre qu'il est possible, en supposant vraies certaines relations sur les paramètres, de construire des formes linéaires répondant essentiellement aux hypothèses des énoncés des critères des parties 5 et 6 ainsi que de construire des formes linéaires mettant en échec les conclusions de ces critères. Cela permet finalement de donner une réciproque aux critères des parties 5 et 6, ce qui démontre qu'ils sont optimaux.

4 Notations

On introduit ici quelques notations préliminaires qui seront utiles dans la suite de cette partie.

Si $(Q_n)_{n\in\mathbb{N}}$ est une suite d'entiers naturels strictement croissante, on définit, pour $Q \geq Q_0$:

$$\Phi(Q) = \max\{k \in \mathbb{N}, Q_k \leq Q\}\,.$$

On a alors clairement les relations

$$Q_{\Phi(Q)} \leq Q < Q_{\Phi(Q)+1}\,.$$

Par ailleurs, étant donné un réseau Λ du dual $(\mathbb{R}^p)^*$ de $\mathbb{R}^p$, on note $\Lambda^\perp$ le réseau dual, c'est-à-dire l'ensemble des vecteurs $x \in \mathbb{R}^p$ tels que $L(x) \in \mathbb{Z}$ pour toute forme linéaire $L \in \Lambda$ (voir par exemple [Cas97] chapitre I.5 page 23). Dans le cas particulier des sections 5 et 6 ci-dessous, $\Lambda_n = \bigoplus_{i=1}^p \delta_{i,n}\mathbb{Z}$ est l'ensemble des formes linéaires $L = \ell_1 X_1 + \cdots + \ell_p X_p$ telles que $\delta_{i,n} | \ell_i$ pour tout $1 \leq i \leq p$. On voit alors que $\Lambda_n^\perp$ est l'ensemble des $(x_1, \ldots, x_p) \in \mathbb{Q}^p$ tels que $\delta_{i,n} x_i \in \mathbb{Z}$ pour tout $i \in \{1, \ldots, p\}$, *i.e.* $\Lambda_n^\perp = 1/\delta_{1,n}\mathbb{Z} \oplus \cdots \oplus 1/\delta_{p,n}\mathbb{Z}$.

Lorsque $(e_1, \ldots, e_p)$ est une base de $\mathbb{R}^p$, on s'intéresse pour $\varepsilon, Q > 0$ au corps convexe suivant (lorsque $\tau_1, \ldots, \tau_{p-1}$ sont fixés) :

$$\mathcal{C}_{Q,\varepsilon} = \left\{ \sum_{i=1}^{p} \lambda_i e_i, |\lambda_i| \leq Q^{\tau_i - \varepsilon} \text{ pour } 1 \leq i \leq p-1 \text{ et } |\lambda_p| \leq Q^{-1-\varepsilon} \right\}$$

qui est symétrique par rapport à l'origine. Dans le cas particulier étudié aux paragraphes 5 et 6 qui consiste à prendre $e_i = (0, \ldots, 0, 1, 0, \ldots, 0, -\xi_i)$ avec le 1 en i-ème position pour $1 \leq i \leq p-1$ et $e_p = (\xi_1, \ldots, \xi_{p-1}, 1)$, avec le réseau Λ_n de la forme $\bigoplus_{i=1}^{p} \delta_{i,n}\mathbb{Z}$, il s'agit essentiellement de

$$\mathcal{C}'_{Q,\varepsilon} = \left\{ \begin{array}{c} (a_1, \ldots, a_p) \in \mathbb{R}^p, |a_i| \leq Q^{\tau_i - \varepsilon} \text{ pour } 1 \leq i \leq p-1 \\ \text{et } |a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| \leq Q^{-1-\varepsilon} \end{array} \right\} .$$

La relation entre $\mathcal{C}_{Q,\varepsilon}$ et $\mathcal{C}'_{Q,\varepsilon}$ est étudiée dans la proposition 7.3. Ces corps convexes ne seront utilisés explicitement qu'à partir de la section 7.

5 Raffinement du critère de Nesterenko

Dans cette partie, on démontre un critère "à la Nesterenko", c'est-à-dire qu'ici les τ_i sont pris 2 à 2 distincts et les estimations sont fortes et de régularité : les formes linéaires sont exactement de taille $Q_n^{-\tau_i+o(1)}$ et de normes $Q_n^{1+o(1)}$.

Théorème 5.1 :
Soient $\xi_1, \ldots, \xi_{p-1}$ des réels. Soient $\tau_1, \ldots, \tau_{p-1} > -1$ des réels 2 à 2 distincts et $(Q_n)_{n\in\mathbb{N}}$ une suite d'entiers strictement croissante telle que $Q_{n+1} = Q_n^{1+o(1)}$. Pour tout $i \in \{1, \ldots, p\}$ et pour tout $n \in \mathbb{N}$, soit $\delta_{i,n} \in \mathbb{N}^*$ tel que $\delta_{i,n}|\delta_{i,n+1}$. Soient maintenant des entiers $\ell_{i,n}$ pour $i = 1, \ldots, p$ et pour tout $n \in \mathbb{N}$, tels que :

(i) $\forall i \in \{1, \ldots, p\}$, $\delta_{i,n}|\ell_{i,n}$;

(ii) $\forall i \in \{1, \ldots, p-1\}$, $|\ell_{i,n} - \ell_{p,n}\xi_i| = Q_n^{-\tau_i+o(1)}$;

(iii) $|\ell_{p,n}| = Q_n^{1+o(1)}$.

Soit $\varepsilon > 0$, Q suffisamment grand par rapport à ε, et $(a_1, \ldots, a_p) \in \mathbb{Q}^p \backslash \{\mathbf{0}\}$ tels que pour tout $i \in \{1, \ldots, p\}$, $\delta_{i,\Phi(Q)}a_i \in \mathbb{Z}$ et pour tout $i \leq p-1$, $|a_i| \leq Q^{\tau_i-\varepsilon}$. Alors on a

$$|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| > Q^{-1-\varepsilon} .$$

Avec les notations de la section 4, la conclusion s'écrit $\Lambda_{\Phi(Q)}^{\perp} \cap \mathcal{C}'_{Q,\varepsilon} = \{\mathbf{0}\}$.

Ce théorème est exactement le théorème 2.1 énoncé dans l'introduction page 13. Il raffine le théorème 1.10 de l'introduction, démontré par Fischler [Fis13] : la principale nouveauté est que les formes linéaires $\ell_{i,n} - \xi_i\ell_{p,n}$ ne tendent plus forcément vers 0, puisque les τ_i peuvent être négatifs. Par ailleurs, les contraintes sur les diviseurs $\delta_{i,n}$ sont affaiblies. La seule contrepartie est qu'on demande aux coefficients $\ell_{i,n}$ d'être, en valeur absolue, égaux à $Q_n^{1+o(1)}$ (il suffit de faire cette hypothèse pour $i = p$, puisque $\tau_i > -1$ pour tout $i \leq p-1$). En pratique, cela ne pose pas de problème (quitte à diminuer Q_n si nécessaire, ce qui améliore la conclusion du théorème), puisqu'on a $|\ell_{p,n}| < |\ell_{p,n+1}| = |\ell_{p,n}|^{1+o(1)}$ dans les applications.

On remarque aussi que sous l'hypothèse supplémentaire $\delta_{i,n} = Q_n^{\gamma_i+o(1)}$ (hypothèse qui sera utile pour les parties 7 et 8), on a nécessairement $a_i = 0$ pour tous les i tel que $\tau_i + \gamma_i < 0$. Le lecteur obtiendra plus de détails sur ce point dans les démonstrations des propositions 7.3, 8.6 et 8.7.

La preuve [Fis13] du théorème 1.10 de l'introduction repose sur le lemme suivant (Lemma 3 de [Fis13]) :

Lemme 5.2 (Fischler) **:**
Soit $M \in \mathcal{M}_p(\mathbb{R})$ une matrice carrée de taille p, dont les coefficients $m_{i,j}$ sont non nuls et vérifient

$$|m_{i',j}m_{i,j'}| \leq \frac{1}{(p+1)!}|m_{i,j}m_{i',j'}| \text{ pour tous } i,j,i',j' \text{ tels que } i < i' \text{ et } j < j' . \qquad (2)$$

Alors M est inversible.

Ce lemme est démontré et utilisé dans [Fis13] (avec une majoration des coeficients de M^{-1} qui est inutile ici) sous l'hypothèse que les coefficients $m_{i,j}$ sont strictement positifs. Mais la preuve de [Fis13] fonctionne plus généralement dès que $m_{i,j} \neq 0$, car elle utilise seulement les majorations (2) sur les valeurs absolues des coefficients.

On commence par reproduire la démonstration du lemme 5.2.

Preuve du lemme 5.2. Considérons $\sigma \in \mathfrak{S}_p \setminus \{\mathrm{Id}\}$. On pose $\kappa_\sigma = \max\{j \leq p, \sigma(j) \neq j\}$. Comme σ est une bijection, on a $\kappa_\sigma \geq 2$. On va alors démontrer, par récurrence sur κ_σ la relation :

$$\prod_{j=1}^{p} |m_{\sigma(j),j}| \leq \frac{1}{(p+1)!}\prod_{j=1}^{p}|m_{j,j}|, \qquad \text{pour } \sigma \neq \mathrm{Id}\,. \qquad (3)$$

Si $\kappa_\sigma = 2$, on a alors $\sigma = \tau_{1,2}$ par définition de κ_σ, où $\tau_{a,b}$ est la transposition échangeant a et b. On a alors

$$\prod_{j=1}^{p} |m_{\sigma(j),j}| = |m_{2,1}||m_{1,2}|\prod_{j=3}^{p}|m_{j,j}|$$

et la relation (2) du lemme 5.2 finit la preuve de la relation (3).

Soit maintenant $\sigma \in \mathfrak{S}_p \setminus \{\mathrm{Id}\}$. Supposons que (3) soit vérifiée pour tout $\sigma' \neq \mathrm{Id}$ tel que $\kappa_{\sigma'} < \kappa_\sigma$.

Par définition de κ_σ, on a $\sigma(j) = j$ pour $\kappa_\sigma + 1 \leq j \leq p$ et $\sigma(\kappa_\sigma) \neq \kappa_\sigma$. Mais si on suppose que $\sigma(\kappa_\sigma) > \kappa_\sigma$, on a alors $\sigma(\sigma(\kappa_\sigma)) = \sigma(\kappa_\sigma)$ ce qui donne $\sigma(\kappa_\sigma) = \kappa_\sigma$ par bijectivité de σ, ce qui est impossible. Ainsi, $\sigma(\kappa_\sigma) < \kappa_\sigma$.

On pose alors $j_0 = \sigma^{-1}(\kappa_\sigma)$. Si on suppose $j_0 > \kappa_\sigma$, alors $j_0 = \sigma(j_0) = \kappa_\sigma$ par définition de j_0 et κ_σ, ce qui aboutit à une contradiction. Et si on suppose $j_0 = \kappa_\sigma$, alors $\kappa_\sigma = \sigma(\kappa_\sigma)$ ce qui est également contradictoire avec la définition de κ_σ. Ainsi, $j_0 < \kappa_\sigma$.

On pose maintenant $\sigma' = \sigma \circ \tau_{j_0,\kappa_\sigma}$. On a donc, pour tout $j \notin \{j_0, \kappa_\sigma\}$, $\sigma'(j) = \sigma(j)$, $\sigma'(j_0) = \sigma(\kappa_\sigma)$ et $\sigma'(\kappa_\sigma) = \kappa_\sigma$, de sorte que $\sigma'(j) = j$ pour $\kappa_\sigma \leq j \leq p$. D'où $\kappa_{\sigma'} < \kappa_\sigma$. On peut donc appliquer (3) :

$$\prod_{j=1}^{p} |m_{\sigma'(j),j}| = |m_{\sigma(\kappa_\sigma),j_0}||m_{\kappa_\sigma,\kappa_\sigma}| \prod_{\substack{1\leq j\leq p\\ j\neq j_0,\kappa_\sigma}} |m_{\sigma(j),j}| \leq \prod_{j=1}^{p} |m_{j,j}| \, . \tag{4}$$

D'autre part, la relation (2) appliquée à $j_0 < \kappa_\sigma$ et $\sigma(\kappa_\sigma) < \kappa_\sigma$, donne

$$|m_{\kappa_\sigma,j_0} m_{\sigma(\kappa_\sigma),\kappa_\sigma}| \leq \frac{1}{(p+1)!} |m_{\sigma(\kappa_\sigma),j_0} m_{\kappa_\sigma,\kappa_\sigma}| \, . \tag{5}$$

La multiplication de (5) par $\prod_{\substack{1\leq j\leq p\\ j\neq j_0,\kappa_\sigma}} |m_{\sigma(j),j}|$ et l'utilisation de (4) achèvent la démonstration de la relation (3) par récurrence.

Finalement, on a :

$$\begin{aligned} |\det(M)| = \left| \sum_{\sigma\in\mathfrak{S}_p} \epsilon_\sigma \prod_{j=1}^{p} m_{\sigma(j),j} \right| &\geq \prod_{j=1}^{p} |m_{j,j}| - \sum_{\substack{\sigma\in\mathfrak{S}_p\\ \sigma\neq \mathrm{Id}}} \prod_{j=1}^{p} |m_{\sigma(j),j}| \\ &\geq \prod_{j=1}^{p} |m_{j,j}| \left(1 - \frac{1}{p+1}\right) > 0 \end{aligned}$$

ce qui termine la preuve du lemme 5.2. □

Pour démontrer le théorème 1.10, Fischler combine ce lemme (appliqué avec $p-1$ au lieu de p) avec le premier théorème de Minkowski sur les corps convexes. Pour démontrer le théorème 5.1, on applique ce lemme à une matrice carrée de taille p (en utilisant notamment l'hypothèse (iii)) et on obtient ainsi p vecteurs linéairement indépendants qui fournissent des approximations simultanées de $\xi_1, \ldots, \xi_{p-1}$ par des nombres rationnels ayant le même dénominateur. On conclut alors la preuve par un argument à la Siegel, sans utiliser la géométrie des nombres. Ceci permet d'éviter les hypothèses sur les diviseurs, présentes dans le théorème 1.10, et de ne pas avoir besoin de supposer $\tau_1, \ldots, \tau_{p-1} > 0$.

Démonstration du théorème 5.1. On raisonne par l'absurde. Quitte à permuter $\xi_1, \ldots, \xi_{p-1}$ qui jouent des rôles symétriques, on peut supposer $\tau_1 > \cdots > \tau_{p-1} > -1$. Soit $\varepsilon > 0$ et Q suffisamment grand par rapport à ε. Soit $P = (a_1, \ldots, a_p) \in \mathbb{Q}^p$, $P \neq \mathbf{0}$ tel que $|a_i| \leq Q^{\tau_i - \varepsilon}$ pour $1 \leq i \leq p-1$, $\delta_{i,\Phi(Q)} a_i \in \mathbb{Z}$ pour tout $i \leq p$ et $|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| \leq Q^{-1-\varepsilon}$.

On prend maintenant $\varepsilon_1 > 0$ assez petit pour que :

$$\begin{cases} \tau_i \left(1 - (1+\varepsilon_1)^{p-1}\right) < \varepsilon \quad \text{pour tout } i \in \{1, \ldots, p-1\} \\ (1+\varepsilon_1)^{p-1} < 1 + \varepsilon \end{cases} \tag{6}$$

On définit $\varphi : \mathbb{N}^* \to \mathbb{N}^*$ par $\varphi(n) - 1 = \Phi(Q_n^{1+\varepsilon_1}) = \max\{k \in \mathbb{N}, Q_k \leq Q_n^{1+\varepsilon_1}\}$. Cette définition a bien un sens car on a évidemment $Q_n \leq Q_n^{1+\varepsilon_1}$, donc l'ensemble est non vide et admet un maximum car $Q_k \xrightarrow[k\to\infty]{} \infty$. De plus, cette remarque nous donne également $\varphi(n) \geq n+1$, d'où $\lim_{n\to\infty} \varphi(n) = +\infty$. Ainsi, $Q_{\varphi(n)} = Q_{\varphi(n)-1}^{1+o(1)}$. Mais d'autre part, on a, par définition de φ que

$$Q_{\varphi(n)-1} \leq Q_n^{1+\varepsilon_1} < Q_{\varphi(n)} = Q_{\varphi(n)-1}^{1+o(1)}\,.$$

On en déduit donc

$$Q_{\varphi(n)} = Q_n^{1+\varepsilon_1+o(1)}\,, \tag{7}$$

de sorte que

$$Q_{\varphi_i(n)} = Q_{\varphi_{i-1}(n)}^{1+\varepsilon_1+o(1)} = Q_n^{(1+\varepsilon_1)^i+o(1)}\,, \tag{7 bis}$$

où φ_i est la i-ème itérée de φ (*i.e.*, $\varphi_i = \underbrace{\varphi \circ \cdots \circ \varphi}_{i \text{ fois}}$), avec $\varphi_0(n) = n$ et $\varphi_1(n) = \varphi(n)$.

On choisit maintenant $n = \Phi(Q)$. Par définition de $\Phi(Q)$, on a donc $Q_n \leq Q < Q_{n+1} \leq Q_{\varphi(n)}$. On remarque alors que $n \xrightarrow[Q\to\infty]{} +\infty$ et donc $o(1)$ est une suite dépendante de n ou Q indifféremment et qui tend vers 0 quand n ou Q tend vers $+\infty$. On pourra donc voir les suites $o(1)$ comme des suites dépendantes de n ou Q sans distinction ; on notera simplement $o(1)$. De plus, l'encadrement précédent nous donne aussi $Q_n = Q^{1+o(1)}$.

On considère maintenant la matrice $p \times p$:

$$M_n = \begin{pmatrix} \ell_{1,n} - \ell_{p,n}\xi_1 & \cdots & \ell_{1,\varphi_{p-1}(n)} - \ell_{p,\varphi_{p-1}(n)}\xi_1 \\ \vdots & \ddots & \vdots \\ \ell_{p-1,n} - \ell_{p,n}\xi_{p-1} & \cdots & \ell_{p-1,\varphi_{p-1}(n)} - \ell_{p,\varphi_{p-1}(n)}\xi_{p-1} \\ \ell_{p,n} & \cdots & \ell_{p,\varphi_{p-1}(n)} \end{pmatrix}.$$

Notons $m_{i,j}$ le coefficient en position (i,j) de cette matrice. Si $1 \leq i \leq p-1$, on a $m_{i,j} = \ell_{i,\varphi_{j-1}(n)} - \ell_{p,\varphi_{j-1}(n)}\xi_i$ pour tout j, donc

$$|m_{i,j}| = Q_{\varphi_{j-1}(n)}^{-\tau_i+o(1)} = Q_n^{-(1+\varepsilon_1)^{j-1}\tau_i+o(1)} \tag{8}$$

en utilisant l'hypothèse (ii), la relation (7 bis) et le fait que $\varphi_k(n) > n$. De même, pour $i = p$, l'hypothèse (iii) et la relation (7 bis) de nouveau donnent

$$|m_{p,j}| = |\ell_{p,\varphi_{j-1}(n)}| = Q_{\varphi_{j-1}(n)}^{1+o(1)} = Q_n^{-(1+\varepsilon_1)^{j-1}\tau_p+o(1)}$$

en posant $\tau_p = -1$. La relation (8) est ainsi encore vraie pour $i = p$ et on a $\tau_1 > \cdots > \tau_{p-1} > \tau_p = -1$.

Si n est assez grand, on a donc $m_{i,j} \neq 0$ pour tous i,j. En outre, pour $i,j,i',j' \in \{1,\ldots,p\}$ tels que $i < i'$ et $j < j'$, on a

$$\begin{aligned} \log\left|\frac{m_{i',j}m_{i,j'}}{m_{i,j}m_{i',j'}}\right| &= -\Big((1+\varepsilon_1)^{j-1}(\tau_{i'}-\tau_i) + (1+\varepsilon_1)^{j'-1}(\tau_i-\tau_{i'}) + o(1)\Big)\log Q_n \\ &= -(\tau_i-\tau_{i'})\Big((1+\varepsilon_1)^{j'-1} - (1+\varepsilon_1)^{j-1} + o(1)\Big)\log Q_n\,. \end{aligned}$$

Or $\tau_i > \tau_{i'}$ et $(1+\varepsilon_1)^{j'-1} > (1+\varepsilon_1)^{j-1}$, donc cette dernière quantité tend vers $-\infty$ quand $n \to +\infty$: pour n assez grand, elle est inférieure à $-\log((p+1)!)$. On peut donc appliquer le lemme 5.2 pour n suffisamment grand et la matrice M_n est alors inversible.

Donc, les lignes de cette matrice sont linéairement indépendantes et toutes leurs combinaisons linéaires non triviales sont non nulles. En particulier, il existe $k \in \{0, \dots, p-1\}$ tel que le réel $A_k = \sum_{i=1}^{p-1} a_i(\ell_{i,\varphi_k(n)} - \xi_i \ell_{p,\varphi_k(n)}) + \left(\sum_{i=1}^{p-1} a_i\xi_i + a_p\right) \ell_{p,\varphi_k(n)}$ soit non nul. En effet, si tous les coefficients de cette combinaison linéaire étaient nuls, on aurait alors $P = (a_1, \dots, a_p) = (0, \dots, 0)$ ce qui est impossible par choix de P. On a donc :

$$\begin{aligned} A_k &= \sum_{i=1}^{p-1} a_i(\ell_{i,\varphi_k(n)} - \xi_i \ell_{p,\varphi_k(n)}) + \left(\sum_{i=1}^{p-1} a_i\xi_i + a_p\right) \ell_{p,\varphi_k(n)} \\ &= \sum_{i=1}^{p} a_i \ell_{i,\varphi_k(n)} \neq 0\,. \end{aligned}$$

Mais on sait par hypothèse que pour tout $i \in \{1, \dots, p\}$, on a $\delta_{i,n} a_i \in \mathbb{Z}$, $\delta_{i,m} \big| \delta_{i,m+1}$ et $\delta_{i,m} \big| \ell_{i,m}$ pour tout $m \in \mathbb{N}$. On en déduit donc pour tout $i \in \{1, \dots, p\}$:

$$a_i \ell_{i,\varphi_k(n)} = \delta_{i,n} a_i \, \frac{\ell_{i,\varphi_k(n)}}{\delta_{i,\varphi_k(n)}} \, \frac{\delta_{i,\varphi_k(n)}}{\delta_{i,n}} \in \mathbb{Z}$$

car $n < \varphi_k(n)$. Donc A_k est un entier non nul.

Mais d'autre part, en utilisant les hypothèses (ii), (iii) et la relation (7 bis), on a les majorations suivantes :

$$\begin{aligned} |A_k| &\leq \sum_{i=1}^{p-1} |a_i| |\ell_{i,\varphi_k(n)} - \xi_i \ell_{p,\varphi_k(n)}| + \left|\sum_{i=1}^{p-1} a_i\xi_i + a_p\right| |\ell_{p,\varphi_k(n)}| \\ &\leq \sum_{i=1}^{p-1} Q^{\tau_i - \varepsilon} Q_{\varphi_k(n)}^{-\tau_i + o(1)} + Q^{-1-\varepsilon} Q_{\varphi_k(n)}^{1+o(1)} \\ &= \sum_{i=1}^{p-1} Q^{\tau_i - \varepsilon} Q_n^{-\tau_i(1+\varepsilon_1)^k + o(1)} + Q^{-1-\varepsilon} Q_n^{(1+\varepsilon_1)^k + o(1)} \\ &= \sum_{i=1}^{p-1} Q^{\tau_i(1-(1+\varepsilon_1)^k) - \varepsilon + o(1)} + Q^{(1+\varepsilon_1)^k - 1 - \varepsilon + o(1)}\,. \end{aligned}$$

Le choix de ε_1 nous permet d'avoir chacun des termes de la somme $< 1/p$ si Q est assez grand (voir les conditions (6)), ce qui aboutit à une contradiction.

$\square$

On remarque que dans la première condition sur ε_1 dans (6), seuls les $i \in \{1, \dots, p-1\}$ tels que $\tau_i < 0$ interviennent réellement. En effet, clairement, si $\tau_i \geq 0$, l'inégalité $\tau_i(1-(1+\varepsilon_1)^{p-1}) < \varepsilon$ est vérifiée quelque soit $\varepsilon_1 > 0$.

6 Critère "à la Siegel"

Dans cette partie, on démontre un critère "à la Siegel" : les hypothèses sont plus faibles et on ne requiert plus que les τ_i soient distincts. En contrepartie, on demande un déterminant non nul, c'est-à-dire d'avoir p approximations des réels ξ_i linéairement indépendantes.

Théorème 6.1 :
Soient $\xi_1, \ldots, \xi_{p-1}$ des réels quelconques, et $\tau_1, \ldots, \tau_{p-1}$ des réels > -1. Soit $(Q_n)_{n\geq 1}$ une suite strictement croissante d'entiers positifs telle que $Q_{n+1} = Q_n^{1+o(1)}$. On suppose également que pour tout $n \in \mathbb{N}$ et pour tout $i \in \{1, \ldots, p\}$, il existe $\delta_{i,n} \in \mathbb{N}^*$ et $\ell_{i,n} \in \mathbb{Z}$ tels que $\delta_{i,n}$ divise $\delta_{i,n+1}$ et :

(i) $\forall i \in \{1, \ldots, p\}$, $\delta_{i,n} | \ell_{i,n}$;

(ii) $\forall i \in \{1, \ldots, p-1\}$, $|\ell_{i,n} - \xi_i \ell_{p,n}| \leq Q_n^{-\tau_i + o(1)}$;

(iii) $|\ell_{p,n}| \leq Q_n^{1+o(1)}$;

(iv) il existe un entier n_1 tel que, pour tout $n \geq n_1$, il existe des réels $\alpha_0(n)$, $\alpha_1(n)$, $\ldots$, $\alpha_{p-1}(n)$ avec $\alpha_0(n) \neq 0$, tels que pour tout $i \in \{1, \ldots, p\}$ on ait :
$$\ell_{i,n+p} = \sum_{j=0}^{p-1} \alpha_j(n) \ell_{i,n+j} .$$

On suppose enfin que si l'on note Δ_n la matrice suivante de taille $p \times p$:
$$\Delta_n = \begin{pmatrix} \ell_{1,n} & \cdots & \ell_{1,n+p-1} \\ \vdots & \ddots & \vdots \\ \ell_{p,n} & \cdots & \ell_{p,n+p-1} \end{pmatrix} ,$$
il existe un certain $n_2 \geq n_1$ tel que $\det(\Delta_{n_2}) \neq 0$.
Soit $\varepsilon > 0$ et Q suffisamment grand en fonction de ε et soit $(a_1, \ldots, a_p) \in \mathbb{Q}^p \setminus \{\mathbf{0}\}$, tel que $\delta_{i,\Phi(Q)} a_i \in \mathbb{Z}$ pour $1 \leq i \leq p$ et $|a_i| \leq Q^{\tau_i - \varepsilon}$ pour $1 \leq i \leq p-1$. Alors on a
$$|a_1 \xi_i + \cdots + a_{p-1} \xi_{p-1} + a_p| > Q^{-1-\varepsilon} .$$

Avec les notations de la section 4, la conclusion de ce théorème s'écrit $\Lambda^{\perp}_{\Phi(Q)} \cap \mathcal{C}'_{Q,\varepsilon} = \{\mathbf{0}\}$.

Ce théorème est exactement le théorème 2.2 énoncé dans l'introduction (page 15). Notons qu'à la place de (iv) et de la non-nullité de $\det(\Delta_{n_2})$, en notant $L_n = (\ell_{1,n}, \ldots, \ell_{p,n})$, on peut supposer que parmi k formes linéaires consécutives $L_n, \ldots, L_{n+k-1}$, il y en a toujours p qui sont linéairement indépendantes (où $k \geq p$ est un entier fixé). En effet, on peut alors noter φ l'extraction telle que $\varphi(n+1) = \min\left\{m > \varphi(n), L_m \notin \mathrm{Vect}(L_{\varphi(n-p+2)}, \ldots, L_{\varphi(n)})\right\}$, et on a $\varphi(n+1) - \varphi(n) \leq k-1$ qui est une constante, de sorte que $Q_{\varphi(n+1)} = Q_{\varphi(n)}^{1+o(1)}$. Comme $L_{\varphi(n)}, \ldots, L_{\varphi(n+p-1)}$ est une base de $(\mathbb{R}^p)^*$, on peut écrire $L_{\varphi(n+p)} = \sum_{j=0}^{p-1} \alpha_j(n) L_{\varphi(n+j)}$

et on a nécessairement $\alpha_0(n) \neq 0$, car sinon $L_{\varphi(n+p)} \in \mathrm{Vect}\big(L_{\varphi(n+1)}, \dots, L_{\varphi(n+p-1)}\big)$ ce qui n'est pas possible par choix de φ. Et le théorème s'applique donc à la suite extraite grâce à φ.

Il est nécessaire en tout cas, d'une façon ou d'une autre, de supposer qu'on dispose de p formes linéaires linéairement indépendantes. Cela permet d'assurer que la combinaison linéaire $\sum_{i=1}^{p} a_i L_{n+i}$ n'est pas nulle, point essentiel de la preuve.

Démonstration. On raisonne, ici encore, par l'absurde. Soit $\varepsilon > 0$ et Q assez grand pour avoir $\Phi(Q) \geq n_2$. Soit également $P = (a_1, \dots, a_p) \neq \mathbf{0}$ avec $|a_i| \leq Q^{\tau_i - \varepsilon}$ pour $1 \leq i \leq p-1$, $a_i \delta_{i,\Phi(Q)} \in \mathbb{Z}$ pour $1 \leq i \leq p$ et $|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| \leq Q^{-1-\varepsilon}$.

De même que dans la démonstration du théorème 5.1, on a $n \xrightarrow[Q \to +\infty]{} +\infty$ si $n = \Phi(Q)$, de sorte que $Q_n = Q^{1+o(1)}$ et les suites $o(1)$ sont des suites tendant vers 0 quand n ou Q tendent vers $+\infty$ indifféremment.

On remarque que pour tout $n \geq n_1$:

$$
\begin{aligned}
\mathrm{rg}(\Delta_{n+1}) &= \mathrm{rg}\begin{pmatrix} \ell_{1,n+1} & \dots & \ell_{1,n+p} \\ \vdots & \ddots & \vdots \\ \ell_{p,n+1} & \dots & \ell_{p,n+p} \end{pmatrix} \\
&= \mathrm{rg}\begin{pmatrix} \ell_{1,n+1} & \dots & \ell_{1,n+p-1} & \sum_{j=0}^{p-1} \alpha_j(n)\ell_{1,n+j} \\ \vdots & \ddots & \vdots & \vdots \\ \ell_{p,n+1} & \dots & \ell_{p,n+p-1} & \sum_{j=0}^{p-1} \alpha_j(n)\ell_{p,n+j} \end{pmatrix} \\
&= \mathrm{rg}\begin{pmatrix} \ell_{1,n+1} & \dots & \ell_{1,n+p-1} & \alpha_0(n)\ell_{1,n} \\ \vdots & \ddots & \vdots & \vdots \\ \ell_{p,n+1} & \dots & \ell_{p,n+p-1} & \alpha_0(n)\ell_{p,n} \end{pmatrix} \\
&= \mathrm{rg}(\Delta_n)
\end{aligned}
$$

car $\alpha_0(n) \neq 0$.

Donc, à partir du rang n_1, tous les Δ_n sont de même rang p, puisque $\det(\Delta_{n_2}) \neq 0$. Donc les lignes de la matrice Δ_n sont linéairement indépendantes, dès que $n \geq n_2$: toute combinaison linéaire non triviale de ces lignes est non nulle. On applique ce résultat à $n = \Phi(Q)$ avec $n \geq n_2$ par hypothèse. Alors il existe $k \in \{0, \dots, p-1\}$ tel que la k-ème coordonnée du vecteur obtenu par combinaison linéaire des lignes de la matrice Δ_n avec $a_1, \dots, a_p$ comme coefficient n'est pas nul, *i.e.* $A_k = \sum_{i=1}^{p} a_i \ell_{i,n+k} \neq 0$.

On sait d'autre part, par hypothèse, que $\delta_{i,n} \big| \delta_{i,n+k} \big| \ell_{i,n+k}$ et $\delta_{i,n} a_i \in \mathbb{Z}$ pour tout $i = 1, \dots, p$. Donc $A_k = \sum_{i=1}^{p} a_i \ell_{i,n+k} = \sum_{i=1}^{p} (\delta_{i,n} a_i) \frac{\ell_{i,n+k}}{\delta_{i,n}} \in \mathbb{Z} \backslash \{0\}$.

De plus, on a les égalités :

$$
\begin{aligned}
A_k &= \sum_{i=1}^{p} a_i \ell_{i,n+k} \\
&= \sum_{i=1}^{p-1} a_i(\ell_{i,n+k} - \xi_i \ell_{p,n+k}) + \left(\sum_{i=1}^{p-1} a_i \xi_i + a_p\right) \ell_{p,n+k}\,.
\end{aligned}
$$

Et par hypothèse, on a

$$\begin{aligned}
\left| \left(\sum_{i=1}^{p-1} a_i \xi_i + a_p \right) \ell_{p,n+k} \right| &\leq Q^{-1-\varepsilon} Q_{n+k}^{1+o(1)} \\
&= Q^{-1-\varepsilon} Q_n^{1+o(1)} && \text{par } k \text{ itérations de } Q_{n+1} = Q_n^{1+o(1)} \\
&= Q^{-1-\varepsilon} Q^{1+o(1)} = Q^{-\varepsilon+o(1)} ,
\end{aligned}$$

et aussi

$$\begin{aligned}
\left| \sum_{i=1}^{p-1} a_i (\ell_{i,n+k} - \xi_i \ell_{p,n+k}) \right| &\leq \sum_{i=1}^{p-1} Q^{\tau_i - \varepsilon} Q_{n+k}^{-\tau_i + o(1)} \\
&= \sum_{i=1}^{p-1} Q^{\tau_i - \varepsilon} Q_n^{-\tau_i + o(1)} && \text{comme précédemment} \\
&= \sum_{i=1}^{p-1} Q^{\tau_i - \varepsilon} Q^{-\tau_i + o(1)} = Q^{-\varepsilon + o(1)} ,
\end{aligned}$$

de sorte qu'on ait les majorations suivantes :

$$\begin{aligned}
|A_k| &\leq \left| \sum_{i=1}^{p-1} a_i (\ell_{i,n+k} - \xi_i \ell_{p,n+k}) \right| + \left| \left(\sum_{i=1}^{p} a_i \xi_i \right) \ell_{p,n+k} \right| \\
&\leq Q^{-\varepsilon+o(1)} + Q^{-\varepsilon+o(1)} \leq Q^{-\varepsilon/2} < 1
\end{aligned}$$

si Q est assez grand.

Donc A_k est un entier non nul de valeur absolue < 1 si Q est assez grand, ce qui est absurde. □

7 Généralisations en termes de corps convexes

7.1 Critères plus généraux

Dans ce paragraphe, on généralise les résultats obtenus ci-dessus (§ 5 et 6) : on utilise une base quelconque $(e_1, \dots, e_p)$ de $\mathbb{R}^p$ (comme dans [Fis13], au lieu d'une base particulière provenant de nombres réels $\xi_1, \dots, \xi_{p-1}$), et des réseaux Λ_n qui ne sont plus forcément de la forme $\bigoplus_{i=1}^p \delta_{i,n}\mathbb{Z}$. On note $\|.\|$ une norme quelconque sur l'espace vectoriel $(\mathbb{R}^p)^*$ des formes linéaires sur $\mathbb{R}^p$.

On utilise les notations introduites en début de partie (voir § 4).

Pour commencer, la généralisation du théorème 5.1 (critère "à la Nesterenko") est la suivante.

Proposition 7.1 :
Soit $(e_1, \dots, e_p)$ une base de $\mathbb{R}^p$ et $\|.\|$ une norme sur $(\mathbb{R}^p)^*$.
Soient $\tau_1, \dots, \tau_{p-1}$ des réels > -1 et 2 à 2 distincts. Soit $(Q_n)_{n\in\mathbb{N}}$ une suite d'entiers strictement croissante telle que $Q_{n+1} = Q_n^{1+o(1)}$. Soit $(\Lambda_n)_{n\in\mathbb{N}}$ une suite de réseaux de $(\mathbb{R}^p)^*$ telle que $\Lambda_{n+1} \subset \Lambda_n$. On considère une suite $(L_n)_{n\in\mathbb{N}}$ de formes linéaires telle que

1. $\forall n \in \mathbb{N}$, $L_n \in \Lambda_n$;
2. $|L_n(e_i)| = Q_n^{-\tau_i+o(1)}$, pour $1 \le i \le p-1$;
3. $\|L_n\| = Q_n^{1+o(1)}$.

Alors pour tout $\varepsilon > 0$ et pour tout Q assez grand en termes de ε, on a

$$\Lambda_{\Phi(Q)}^{\perp} \cap \mathcal{C}_{Q,\varepsilon} = \{\mathbf{0}\}\,.$$

On rappelle que, par définition, $\Lambda_n^{\perp}$ est l'ensemble des vecteurs $x \in \mathbb{R}^p$ tels que $L(x) \in \mathbb{Z}$ pour toute forme linéaire $L \in \Lambda_n$ (voir par exemple [Cas97] chapitre I.5 page 23). Dans le cas particulier du théorème 5.1, $\Lambda_n = \bigoplus_{i=1}^{p} \delta_{i,n}\mathbb{Z}$ est l'ensemble des formes linéaires $L = \ell_1 X_1 + \cdots + \ell_p X_p$ telles que $\delta_{i,n} | \ell_i$ pour tout $1 \le i \le p$. On voit alors que $\Lambda_n^{\perp}$ est l'ensemble des $(x_1, \dots, x_p) \in \mathbb{Q}^p$ tels que $\delta_{i,n} x_i \in \mathbb{Z}$ pour tout $i \in \{1, \dots, p\}$.

Lorsque $\xi_1, \dots, \xi_{p-1}$ sont des réels et qu'on pose $e_i = (0, \dots, 0, 1, 0, \dots, 0, -\xi_i)$ où le 1 est en i-ème position pour $1 \le i \le p-1$, et $e_p = (\xi_1, \dots, \xi_{p-1}, 1)$, on obtient, comme cas particulier de cette proposition, un résultat équivalent au théorème 5.1, en prenant aussi $\Lambda_n = \bigoplus_{i=1}^{p} \delta_{i,n}\mathbb{Z}$ (voir [Dau14] pour plus de détails sur la façon dont le théorème 5.1 peut se déduire de **??**). Cette équivalence sera démontrée ci-dessous (proposition 7.3).

L'intérêt de la proposition 7.1 est qu'elle est plus générale que le théorème 5.1 : la base $(e_1, \dots, e_p)$ de $\mathbb{R}^p$ est quelconque et les réseaux Λ_n aussi (et plus nécessairement de la forme $\bigoplus_{i=1}^{p} \delta_{i,n}\mathbb{Z}$).

La preuve de la proposition 7.1 est similaire à celle du théorème 5.1 : le contexte plus général n'introduit pas de difficulté réelle.

Démonstration. Tout d'abord, notons $|||.|||$ la norme sur $(\mathbb{R}^p)^*$ subordonnée à la norme N sur $\mathbb{R}^p$ définie par $N(\mu_1 e_1 + \cdots + \mu_p e_p) = \max_{1\le i\le p} |\mu_i|$. Pour tout $n \in \mathbb{N}$, il existe $x = \mu_1 e_1 + \cdots + \mu_p e_p$ tel que $N(x) = 1$ et

$$|||L_n||| = |L_n(x)| \le \sum_{i=1}^{p} |\mu_i L_n(e_i)| \le |L_n(e_p)| + \sum_{i=1}^{p-1} Q_n^{-\tau_i+o(1)}\,.$$

Comme $\tau_i > -1$ pour tout $1 \le i \le p-1$, l'hypothèse 3. montre (puisque toutes les normes sont équivalentes sur $(\mathbb{R}^p)^*$) que $|L_n(e_p)| \ge Q_n^{1+o(1)}$. Comme l'inégalité dans l'autre sens découle immédiatement de l'hypothèse 3., on a égalité.

Soit $\varepsilon > 0$ et Q suffisamment grand par rapport à ε. Soit $\varepsilon_1 > 0$ tel que

$$\begin{cases} \tau_i \left(1 - (1+\varepsilon_1)^{p-1}\right) < \frac{\varepsilon}{2} & \text{pour tout } i \in \{1, \dots, p-1\} \\ (\varepsilon_1 + 1)^{p-1} - 1 < \frac{\varepsilon}{2} \end{cases} . \tag{9}$$

Comme dans la démonstration du théorème 5.1, seuls les $i \in \{1, \dots, p-1\}$ tels que $\tau_i < 0$ fournissent une contrainte sur ε_1.

On pose aussi $\varphi(n) - 1 = \Phi(Q_n^{1+\varepsilon_1}) = \max\{k \in \mathbb{N},\ Q_k \leq Q_n^{1+\varepsilon_1}\}$. De même que précédemment, on a $Q_{\varphi(n)} = Q_n^{1+\varepsilon_1+o(1)}$. On pose alors $n = \Phi(Q)$ et le même raisonnement que plus haut fournit aussi $Q_n = Q^{1+o(1)}$.

En posant $M_n = (L_{\varphi_{i-1}(n)}(e_j))_{1 \leq i,j \leq p} \in \mathcal{M}_p(\mathbb{R})$, on voit, comme plus haut, en utilisant la relation $|L_n(e_p)| = Q_n^{1+o(1)}$, que le Lemma 3 de [Fis13] (rappelé en lemme 5.2 page 27) s'applique. Ainsi, la matrice M_n est inversible. Donc, les images de la base $e_1, \dots, e_p$ par les formes linéaires $L_n, \dots, L_{\varphi_{p-1}(n)}$ sont linéairement indépendantes, et donc les formes linéaires aussi, si n (et donc Q) est assez grand.

On va montrer que $\mathcal{C}_{Q,\varepsilon} \cap \Lambda_{\Phi(Q)}^{\perp} = \{\mathbf{0}\}$ par l'absurde. On suppose qu'il existe $P \in \mathcal{C}_{Q,\varepsilon} \cap \Lambda_{\Phi(Q)}^{\perp}$ non nul. Comme Q est assez grand, $L_n, \dots, L_{\varphi_{p-1}(n)}$ sont linéairement indépendantes et forment donc une base de l'espace dual de $\mathbb{R}^p$. Comme $P \neq \mathbf{0}$, il existe $k \in \{0, \dots, p-1\}$ tel que $L_{\varphi_k(n)}(P) \neq 0$.

Par hypothèse sur la suite de réseaux, on a $\Lambda_n^{\perp} \subset \Lambda_{n+1}^{\perp}$, et donc $\Lambda_n^{\perp} \subset \Lambda_{\varphi_k(n)}^{\perp}$ par itération. Donc, par définition du réseau dual, $L_{\varphi^k(n)}(P)$ est un entier qui est non nul.

Mais en notant $P = \sum_{i=1}^{p} \lambda_i e_i$ comme dans la définition de $\mathcal{C}_{Q,\varepsilon}$, on a :

$$\begin{aligned}
|L_{\varphi_k(n)}(P)| &\leq \sum_{i=1}^{p} |\lambda_i| |L_{\varphi_k(n)}(e_i)| \\
&\leq \sum_{i=1}^{p-1} Q^{\tau_i - \varepsilon} Q_{\varphi_k(n)}^{-\tau_i + o(1)} + Q^{-1-\varepsilon} Q_{\varphi_k(n)}^{1+o(1)} \\
&= \sum_{i=1}^{p-1} Q^{\tau_i - \varepsilon} \left(Q_n^{(1+\varepsilon_1)^k + o(1)}\right)^{-\tau_i + o(1)} + Q^{-1-\varepsilon} \left(Q_n^{(1+\varepsilon_1)^k + o(1)}\right)^{1+o(1)} \\
&\leq \sum_{i=1}^{p-1} Q^{\tau_i(1-(1+\varepsilon_1)^k) - \varepsilon/2} + Q^{(1+\varepsilon_1)^k - 1 - \varepsilon/2} .
\end{aligned}$$

Mais par le choix (9) de ε_1, chacun des exposants est < 0. Comme Q est assez grand, on a $|L_{\varphi_k(n)}(P)| < 1$, ce qui est contradictoire puisque c'est un entier non nul. □

On peut généraliser de façon analogue le théorème 6.1 du § 6. On obtient le résultat suivant "à la Siegel" :

Proposition 7.2 :
Soit $(e_1, \dots, e_p)$ une base de $\mathbb{R}^p$ et $\|.\|$ une norme sur $(\mathbb{R}^p)^*$.
Soient $\tau_1, \dots, \tau_{p-1}$ des réels > -1 et $(Q_n)_{n\in\mathbb{N}}$ une suite d'entiers strictement croissante telle que $Q_{n+1} = Q_n^{1+o(1)}$. Soit $(\Lambda_n)_{n\in\mathbb{N}}$ une suite de réseaux de $(\mathbb{R}^p)^*$ telle que $\Lambda_{n+1} \subset \Lambda_n$.
On considère une suite $(L_n)_{n\in\mathbb{N}}$ de formes linéaires telle que

1. pour tout $n \in \mathbb{N}$, $L_n \in \Lambda_n$;
2. $|L_n(e_i)| \leq Q_n^{-\tau_i+o(1)}$, pour $1 \leq i \leq p-1$;
3. $\|L_n\| \leq Q_n^{1+o(1)}$;
4. la suite de formes linéaires vérifie une relation de récurrence : il existe $n_1 \geq 0$ tel que pour tout $n \geq n_1$, il existe des nombres réels $\alpha_0(n), \dots, \alpha_{p-1}(n)$ tels que $L_{n+p} = \sum_{i=0}^{p-1} \alpha_i(n) L_{n+i}$ avec $\alpha_0(n) \neq 0$.

On suppose enfin qu'il existe $n_2 \geq n_1$ tel que $L_{n_2}, \dots, L_{n_2+p-1}$ soient linéairement indépendantes.
Alors pour tout $\varepsilon > 0$, pour tout Q assez grand en termes de ε, on a

$$\Lambda_{\Phi(Q)}^{\perp} \cap \mathcal{C}_{Q,\varepsilon} = \{\mathbf{0}\}.$$

Comme précédemment, les ξ_i considérés dans le théorème 6.1 sont remplacés par une base $(e_1, \dots, e_p)$ générale de $\mathbb{R}^p$ et les réseaux particuliers du théorème 6.1 de la forme $\bigoplus_{i=1}^p \delta_{i,n}\mathbb{Z}$ sont ici remplacés par des réseaux quelconques Λ_n. La proposition 7.3 ci-dessous permet de faire le lien entre le théorème 6.1 et la proposition 7.2 (voir [Dau14]).

De même que pour le théorème 6.1, on pourrait remplacer l'existence de n_1 et n_2 par une autre hypothèse permettant d'obtenir des formes linéaires linéairement indépendantes. Par exemple, on pourrait supposer que pour tout n, la famille $(L_n, L_{n+1}, \dots, L_{n+k-1})$ engendre l'espace vectoriel $(\mathbb{R}^p)^*$ (où k est un entier fixé $\geq p$).

Démonstration. Comme au début de la preuve de la proposition 7.1, on remarque que $|L_n(e_p)| \leq Q_n^{1+o(1)}$. Par ailleurs, la matrice

$$\Delta_n = \begin{pmatrix} L_n(e_1) & \dots & L_{n+p-1}(e_1) \\ \vdots & \ddots & \vdots \\ L_n(e_p) & \dots & L_{n+p-1}(e_p) \end{pmatrix}$$

est inversible pour $n = n_2$, et on a $\mathrm{rg}(\Delta_{n+1}) = \mathrm{rg}(\Delta_n)$ pour tout $n \geq n_1$ (comme dans la preuve du théorème 7.1), donc Δ_n est inversible pour tout $n \geq n_1$.

Soit $\varepsilon > 0$ et Q assez grand en termes de ε. Soit $P \in \Lambda_{\Phi(Q)}^{\perp} \cap \mathcal{C}_{Q,\varepsilon} \setminus \{\mathbf{0}\}$. Du fait de la croissance de la suite de réseaux, on a $P \in \Lambda_m^{\perp}$ pour tout $m \geq \Phi(Q)$, et donc $L_m(P) \in \mathbb{Z}$. En prenant alors $n = \Phi(Q)$ et Q assez grand pour avoir $n \geq n_1$, on a $L_n(P) \in \mathbb{Z}$. Mais comme $\det(\Delta_n) \neq 0$, on sait que toute combinaison linéaire non triviale des lignes de la matrice Δ_n est non nulle. En particulier, il existe $k \in \{0, \dots, p-1\}$ tel que $L_{n+k}(P) \neq 0$.

Remarquons maintenant que comme $k \leq p-1$, on a $Q_{n+k} = Q_n^{1+o(1)}$. Mais le choix de $n = \Phi(Q)$ nous donne alors que $n \xrightarrow[Q\to\infty]{} +\infty$ et $Q_n = Q^{1+o(1)}$.

La fin de la preuve est alors identique à celle de la proposition 7.1 : l'entier non nul $L_{n+k}(P)$ a une valeur absolue < 1, ce qui est contradictoire.

□

On remarquera que la propriété 7.2 précédente peut, en fait, être aussi démontrée en appliquant un résultat de Mahler [Mah39] sur les minimums successifs d'un convexe polaire par rapport à un réseau dual. On rappelle ici quelques propriétés sur les minimums successifs. Tous ces résultats peuvent être trouvés en détails dans [Cas97] Chapitre VIII, par exemple.

Si $\mathcal{K}$ est un compact, convexe, symétrique par rapport à l'origine, non réduit à $\{\mathbf{0}\}$, d'un espace euclidien E de dimension d, si $\lambda\mathcal{K} = \{\lambda x, x \in \mathcal{K}\}$ pour tout $\lambda \in \mathbb{R}$ et Λ est un réseau, on note $\lambda_k(\mathcal{K},\Lambda)$ le minimum des réels $\lambda > 0$ tels que $(\lambda\mathcal{K}) \cap \Lambda$ contienne k vecteurs linéairement indépendants. En particulier, $\lambda_1(\mathcal{K},\Lambda)$ est le plus petit réel $\lambda > 0$ tel que $(\lambda\mathcal{K}) \cap \Lambda \neq \{\mathbf{0}\}$. Parallèlement, $\lambda_d(\mathcal{K},\Lambda)$ est le plus petit réel $\lambda > 0$ tel que $(\lambda\mathcal{K}) \cap \Lambda$ contienne une base de E. On a alors

$$0 < \lambda_1(\mathcal{K},\Lambda) \leq \cdots \leq \lambda_d(\mathcal{K},\Lambda)\,,$$

$$\lambda_k(\mathcal{K},\Lambda) \geq \lambda_k(\mathcal{K}',\Lambda')$$

pour tout $1 \leq k \leq d$ et avec $\mathcal{K}' \subset \mathcal{K}$ et $\Lambda' \subset \Lambda$, et

$$\lambda_k(\mu\mathcal{K},\Lambda) = \mu^{-1}\lambda_k(\mathcal{K},\Lambda)$$

$$\lambda_k(\mathcal{K},\mu\Lambda) = \mu\lambda_k(\mathcal{C},\Lambda)$$

pour tout $1 \leq k \leq d$ et tout $\mu > 0$.

Le résultat de Mahler énonce alors

$$1 \leq \lambda_k(\mathcal{K},\Lambda)\lambda_{d+1-k}(\mathcal{K}^*,\Lambda^\perp) \leq d!^2$$

pour tout $1 \leq k \leq d$ et où $\mathcal{K}^*$ est le convexe polaire associé à $\mathcal{K}$ (c'est-à-dire le compact convexe symétrique par rapport à l'origine composé de tous les vecteurs $x \in E$ tels que pour tout $y \in \mathcal{K}$, $x \cdot y \leq 1$) et $\Lambda^\perp$ est le réseau dual de Λ (voir la section 4 ou la remarque qui suit la proposition 7.1). En fait, la minoration suffit à démontrer la proposition 7.2.

7.2 Lien avec les mesures d'indépendance linéaire restreinte

Pour faire le lien entre les deux résultats ci-dessus et ceux démontrés aux paragraphes 5 et 6, on démontre le résultat suivant.

Proposition 7.3 :
Soient $\xi_1,\dots,\xi_{p-1}$ des réels. On pose $e_i=(0,\dots,0,1,0,\dots,0,-\xi_i)$ où le 1 est en i-ème position et $e_p=(\xi_1,\dots,\xi_{p-1},1)$.
Soit $\tau_1,\dots,\tau_{p-1}$ des réels >-1. On prend aussi pour tout $i\in\{1,\dots,p\}$ et pour tout $n\in\mathbb{N}$, $\delta_{i,n}\in\mathbb{N}^*$ avec $\delta_{i,n}|\delta_{i,n+1}$. On note $\Lambda_n=\delta_{1,n}\mathbb{Z}\oplus\cdots\oplus\delta_{p,n}\mathbb{Z}$ le réseau de $\mathbb{R}^p$ qu'ils définissent. On considère aussi $(Q_n)_{n\in\mathbb{N}}$ une suite d'entiers strictement croissante telle que $Q_{n+1}=Q_n^{1+o(1)}$.
On a alors l'équivalence entre les deux propriétés suivantes :

$$\begin{array}{r}\text{Pour tout } \varepsilon>0 \text{ et tout } Q \text{ suffisamment grand par rapport à } \varepsilon,\\ \text{il n'existe pas de vecteur non nul de } \Lambda_{\Phi(Q)}^{\perp} \text{ qui s'écrive sous la forme}\\ \lambda_1e_1+\cdots+\lambda_{p-1}e_{p-1}+ue_p,\ \text{avec } |\lambda_i|\le Q^{\tau_i-\varepsilon}, |u|\le Q^{-1-\varepsilon}\\ \text{où } \lambda_1,\dots,\lambda_{p-1},u\in\mathbb{R}\end{array}\tag{10}$$

et

$$\begin{array}{r}\text{Pour tout } \varepsilon'>0 \text{ et tout } Q' \text{ suffisamment grand par rapport à } \varepsilon',\\ \text{pour tout } (a_1,\dots,a_p)\in\Lambda_{\Phi(Q')}^{\perp}\setminus\{\mathbf{0}\},\ \text{si } |a_i|\le Q'^{\tau_i-\varepsilon'} \text{ pour tout } i\le p-1,\\ \text{alors } |a_1\xi_1+\cdots+a_{p-1}\xi_{p-1}+a_p|>Q'^{-1-\varepsilon'}.\end{array}\tag{11}$$

Avec les notations de la section 4, la propriété (10) s'écrit $\Lambda_{\Phi(Q)}^{\perp}\cap\mathcal{C}_{Q,\varepsilon}=\{\mathbf{0}\}$: c'est la conclusion des propositions 7.1 et 7.2. Quant à elle, la propriété (11) est la conclusion des critères démontrés aux paragraphes 5 et 6 (théorèmes 5.1 et 6.1) : elle s'écrit $\Lambda_{\Phi(Q)}^{\perp}\cap\mathcal{C}'_{Q',\varepsilon'}=\{\mathbf{0}\}$. On voit donc que les théorèmes 5.1 et 6.1 sont des cas particuliers des propositions 7.1 et 7.2 respectivement.

Démonstration. Commençons par montrer que (11) implique (10) en raisonnant par l'absurde.

Soit $\varepsilon>0$ et Q suffisamment grand par rapport à ε. Soit $P=\lambda_1e_1+\cdots+\lambda_{p-1}e_{p-1}+ue_p$ avec $|\lambda_i|\le Q^{\tau_i-\varepsilon}$ et $|u|\le Q^{-1-\varepsilon}$. On suppose $P\neq\mathbf{0}$ et $P\in\Lambda_{\Phi(Q)}^{\perp}$.

On a $P=(a_1,\dots,a_p)=(\lambda_1+u\xi_1,\dots,\lambda_{p-1}+u\xi_{p-1},u-\sum_{k=1}^{p-1}\lambda_k\xi_k)$ avec $\delta_{i,\Phi(Q)}a_i\in\mathbb{Z}$. Pour tout $i\in\{1,\dots,p-1\}$, on a donc

$$|a_i|\le\left(Q^{\tau_i-\varepsilon}+|\xi_i|Q^{-1-\varepsilon}\right)\le(1+|\xi_i|)Q^{\tau_i-\varepsilon}\ \text{puisque } \tau_i>-1\,.\tag{12}$$

Fixons maintenant $\varepsilon'=\frac{\varepsilon}{2}$ et posons

$$Q'=\left(\frac{1}{2+\sum_{1\le i\le p-1}\xi_i^2}Q^{1+\varepsilon}\right)^{\frac{1}{1+\varepsilon'}}.$$

On peut avoir Q' aussi grand que l'on veut par rapport à ε' en prenant Q assez grand, puisque $Q'\xrightarrow[Q\to+\infty]{}+\infty$.

On remarque que les inégalités suivantes sont équivalentes :

$$\begin{aligned} & \tau_i - \varepsilon - (\tau_i - \varepsilon')\frac{1+\varepsilon}{1+\varepsilon'} < 0 \\ \iff & (\tau_i - \varepsilon)(1+\varepsilon') - (\tau_i - \varepsilon')(1+\varepsilon) < 0 \\ \iff & (\tau_i + 1)(\varepsilon' - \varepsilon) < 0\,. \end{aligned}$$

Or $\tau_i > -1$ et $\varepsilon' = \varepsilon/2$, donc ces inégalités sont vérifiées ; par définition de Q', on a donc pour tout $i \in \{1, \dots, p-1\}$:

$$(1+|\xi_i|)Q^{\tau_i - \varepsilon} < \left(\frac{1}{2 + \sum_{k=1}^{p-1} \xi_k^2}\right)^{\frac{\tau_i - \varepsilon'}{1+\varepsilon'}} Q^{(\tau_i - \varepsilon')\frac{1+\varepsilon}{1+\varepsilon'}} = Q'^{\tau_i - \varepsilon'}$$

si Q est assez grand en fonction de ε et puisque $\xi_1, \dots, \xi_{p-1}$ sont fixés. La majoration (12) fournit donc $|a_i| < Q'^{\tau_i - \varepsilon'}$ pour tout $i \in \{1, \dots, p-1\}$.

Enfin, comme $\varepsilon' = \frac{\varepsilon}{2}$, on a $Q' > Q$ (par définition de Q', en supposant Q assez grand par rapport à ε), donc $\Phi(Q') \geq \Phi(Q)$, ce qui donne $\Lambda^{\perp}_{\Phi(Q)} \subset \Lambda^{\perp}_{\Phi(Q')}$ puisque $\delta_{i,\Phi(Q)} | \delta_{i,\Phi(Q')}$ pour tout $i \in \{1, \dots, p\}$. Ainsi, $P = (a_1, \dots, a_p) \in \Lambda^{\perp}_{\Phi(Q')} \setminus \{\mathbf{0}\}$ et on peut appliquer (11), ce qui donne :

$$|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| > Q'^{-1-\varepsilon'}\,.$$

Mais :

$$\sum_{i=1}^{p-1} a_i\xi_i = \sum_{i=1}^{p-1} (\lambda_i + u\xi_i)\xi_i = -a_p + u\left(1 + \sum_{i=1}^{p-1} \xi_i^2\right)\,.$$

On a donc $|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| = |u|\left(1 + \sum_{i=1}^{p-1} \xi_i^2\right) \geq Q'^{-1-\varepsilon'}$. Or on a $|u| \leq Q^{-1-\varepsilon}$, ce qui donne :

$$\left(1 + \sum_{i=1}^{p-1} \xi_i^2\right) Q^{-1-\varepsilon} \geq |u|\left(1 + \sum_{i=1}^{p-1} \xi_i^2\right) \geq Q'^{-1-\varepsilon'}\,.$$

Et la définition de Q' donne alors :

$$1 \leq \left(1 + \sum_{i=1}^{p-1} \xi_i^2\right) Q'^{1+\varepsilon'} Q^{-1-\varepsilon} = \frac{1 + \sum_{i=1}^{p-1} \xi_i^2}{2 + \sum_{i=1}^{p-1} \xi_i^2}$$

ce qui est absurde.

On va montrer maintenant que (10) implique (11) par l'absurde.

Soit $\varepsilon' > 0$ et Q' suffisamment grand par rapport à ε'. On prend $(a_1, \dots, a_p) \in \Lambda^{\perp}_{\Phi(Q')}$ non nul, avec $|a_i| \leq Q'^{\tau_i - \varepsilon'}$ pour $1 \leq i \leq p-1$ et $|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| \leq Q'^{-1-\varepsilon'}$. On va montrer qu'on aboutit à une contradiction.

On pose, pour $i \in \{1, \dots, p-1\}$,

$$\lambda_i = a_i - \xi_i \frac{\sum_{k=1}^{p-1} a_k \xi_k + a_p}{1 + \sum_{k=1}^{p-1} \xi_k^2}$$

et

$$u = \frac{\sum_{1 \le k \le p-1} a_k \xi_k + a_p}{1 + \sum_{1 \le k \le p-1} \xi_k^2}.$$

On obtient alors les relations suivantes pour $i \in \{1, \dots, p-1\}$: $a_i = \lambda_i + u\xi_i$ et $a_p = u - \sum_{1 \le i \le p-1} \lambda_i \xi_i$. On pose $P = \lambda_1 e_1 + \dots + \lambda_{p-1} e_{p-1} + u e_p$, si bien que $P = (a_1, \dots, a_{p-1}, a_p) \in \Lambda^{\perp}_{\Phi(Q')} \setminus \{\mathbf{0}\}$ par construction.

On prend $\varepsilon = \frac{\varepsilon'}{2}$ et on pose

$$Q = \left(\left(1 + \sum_{i=1}^{p-1} \xi_i^2 \right) Q'^{1+\varepsilon'} \right)^{\frac{1}{1+\varepsilon}}.$$

Q peut donc être pris aussi grand que nécessaire par rapport à ε, en choisissant Q' assez grand. On a les majorations suivantes :

$$\begin{aligned} |u| &= \frac{1}{1 + \sum_{1 \le i \le p-1} \xi_i^2} |a_1 \xi_1 + \dots + a_{p-1}\xi_{p-1} + a_p| \\ &\le \frac{1}{1 + \sum_{1 \le i \le p-1} \xi_i^2} Q'^{-1-\varepsilon'} = Q^{-1-\varepsilon}. \end{aligned}$$

Et pour $i \in \{1, \dots, p-1\}$, on a :

$$\begin{aligned} |\lambda_i| &\le |a_i| + \frac{|\xi_i|}{1 + \sum_{k=1}^{p-1} \xi_k^2} \left| \sum_{k=1}^{p-1} a_k \xi_k + a_p \right| \\ &\le Q'^{\tau_i - \varepsilon'} + \frac{|\xi_i|}{1 + \sum_{k=1}^{p-1} \xi_k^2} Q'^{-1-\varepsilon'} \\ &\le \left(1 + \frac{|\xi_i|}{1 + \sum_{k=1}^{p-1} \xi_k^2} \right) Q'^{\tau_i - \varepsilon'} && \text{car } \tau_i > -1 \\ &\le \left(1 + \sum_{k=1}^{p-1} \xi_k^2 \right)^{\frac{\tau_i - \varepsilon'/2}{1+\varepsilon'}} Q'^{\tau_i - \frac{\varepsilon'}{2}} && \text{si } Q' \text{ est assez grand} \\ &= \left(\left(1 + \sum_{k=1}^{p-1} \xi_k^2 \right) Q'^{1+\varepsilon'} \right)^{\frac{\tau_i - \varepsilon'/2}{1+\varepsilon'}} \\ &= Q^{\frac{1+\varepsilon}{1+\varepsilon'} \left(\tau_i - \frac{\varepsilon'}{2} \right)} \le Q^{\tau_i - \varepsilon} \end{aligned}$$

par définition de Q et car $\frac{1+\varepsilon}{1+\varepsilon'} = \frac{1+\varepsilon}{1+2\varepsilon} < 1$ et $\varepsilon'/2 = \varepsilon$.

Finalement, on a $P = \sum_{1\le i\le p-1} \lambda_i e_i + u e_p \neq \mathbf{0}$ dans $\Lambda^{\perp}_{\Phi(Q')}$ avec $|\lambda_i| \le Q^{\tau_i-\varepsilon}$ et $|u| \le Q^{-1-\varepsilon}$. Et comme $\frac{1+\varepsilon'}{1+\varepsilon} > 1$, on a $Q = \left(1+\sum_{i=1}^{p-1}\xi_i^2\right)^{\frac{1}{1+\varepsilon}} Q'^{\frac{1+\varepsilon'}{1+\varepsilon}} > Q'$ si Q est assez grand. Donc, $\Phi(Q) \ge \Phi(Q')$, et pour tout $i \in \{1,\dots,p-1\}$, $\delta_{i,\Phi(Q')}$ divise $\delta_{i,\Phi(Q)}$. Ainsi $\Lambda^{\perp}_{\Phi(Q')} \subset \Lambda^{\perp}_{\Phi(Q)}$ d'où $P \in \Lambda_{\Phi(Q)}$ et on a une contradiction avec (10).

□

8 Optimalité

Dans cette partie, on met à l'épreuve l'optimalité des conclusions obtenues dans les critères des parties précédentes. Compte tenu des hypothèses de comportements asymptotiques (normes, petitesse etc...) on vérifie que ces conclusions sont optimales en construisant une réciproque aux énoncés des théorèmes 5.1 et 6.1. Il ne s'agit cependant pas de réciproques au sens propre du terme à ces énoncés, car les formes linéaires qu'on va construire (en supposant fausses les conclusions des critères, c'est-à-dire les mesures d'indépendance linéaire restreinte) ne satisferont pas (*a priori*) à toutes les hypothèses de ces critères.

Dans le cas le plus simple où on considère un seul nombre et pas de diviseurs, le critère de Nesterenko se réduit au lemme classique suivant.

Lemme 8.1 :
Soit $\xi \in \mathbb{R}\setminus\mathbb{Q}$, et $\alpha,\beta \in \mathbb{R}$ tels que $0<\alpha<1$ et $\beta>1$. Supposons qu'il existe des suites d'entiers $(u_n)_{n\ge1}$ et $(v_n)_{n\ge1}$ tels que

$$\lim_{n\to+\infty} |u_n\xi - v_n|^{1/n} = \alpha \text{ et } \limsup_{n\to+\infty} |u_n|^{1/n} \le \beta\,.$$

Alors $\mu(\xi) \le 1 - \frac{\log\beta}{\log\alpha}$.

Dans ce cas particulier, Fischler et Rivoal ont quasiment démontré [FR10] la réciproque, au sens propre du terme, de ce résultat. Ils obtiennent même des suites $(u_n)_{n\ge1}$ et $(v_n)_{n\ge1}$ satisfaisant à des estimations asymptotiques plus précises que celles des hypothèses du lemme 8.1. Le seul défaut pour que l'énoncé suivant soit vraiment la réciproque du lemme 8.1 est l'inégalité stricte, et non large, sur $\mu(\xi)$ dans l'hypothèse.

Théorème 8.2 (Fischler-Rivoal, 2009) :
Soit $\xi \in \mathbb{R} \setminus \mathbb{Q}$ et $\alpha, \beta \in \mathbb{R}$ tels que $0 < \alpha < 1$ et $\beta > 1$. On suppose $\mu(\xi) < 1 - \frac{\log \beta}{\log \alpha}$. Alors il existe deux suites d'entiers $(u_n)_{n\geq 1}$ et $(v_n)_{n\geq 1}$ telles que

$$\lim_{n\to+\infty} \frac{u_n\xi - v_n}{\alpha^n} = \lim_{n\to+\infty} \frac{u_n}{\beta^n} = 1$$

et par conséquent,

$$\lim_{n\to+\infty} \frac{|u_{n+1}\xi + v_{n+1}|}{|u_n\xi + v_n|} = \lim_{n\to+\infty} |u_n\xi + v_n|^{1/n} = \alpha$$

et

$$\lim_{n\to+\infty} \frac{u_{n+1}}{u_n} = \lim_{n\to+\infty} u_n^{1/n} = \beta\,.$$

Dans [Fis09], Fischler établit également une réciproque presque complète pour l'exposant d'approximation rationnelle restreinte μ_ψ qui inclut des diviseurs (voir la définition 3.2 page 19 dans l'introduction pour la notation μ_ψ et les fonctions $\psi \in \mathcal{E}$).

Théorème 8.3 (Fischler, 2009) :
Soit $\xi \in \mathbb{R} \setminus \mathbb{Q}$, $0 < \alpha < 1$, $\beta > 1$, $\delta \geq 1$ et $(\delta_n)_{n\geq 1}$ une suite d'entiers naturels non nuls telle que δ_n divise δ_{n+1} pour tout $n \geq 1$ et $\delta_n^{1/n} \xrightarrow[n\to+\infty]{} \delta$. Posons $\psi(q) = \delta_n$ où $n = \lfloor \frac{\log q}{\log(\delta/\alpha)} \rfloor$. Alors $\psi \in \mathcal{E}$ et les implications suivantes sont vraies :

- S'il existe deux suites d'entiers $(u_n)_{n\geq 1}$ et $(v_n)_{n\geq 1}$ telles que $u_n \geq 0$, $|u_n\xi - v_n|^{1/n} \xrightarrow[n\to+\infty]{} \alpha$, $u_n^{1/n} \xrightarrow[n\to+\infty]{} \beta$ et δ_n divise u_n pour tout n, alors

$$\mu_\psi(\xi) \leq \frac{\log \beta - \log \alpha}{\log \delta - \log \alpha}\,.$$

- Si l'on a

$$\mu_\psi(\xi) < \frac{\log \beta - \log \alpha}{\log \delta - \log \alpha}\,,$$

  alors il existe deux suites d'entiers $(u_n)_{n\geq 1}$ et $(v_n)_{n\geq 1}$ tel que $u_n \geq 0$, $|u_n\xi - v_n|^{1/n} \xrightarrow[n\to+\infty]{} \alpha$, $u_n^{1/n} \xrightarrow[n\to+\infty]{} \beta$ et δ_n divise u_n pour tout n.

Dans ce résultat, le premier point généralise le lemme 8.1, et le second le théorème 8.2. Il s'agit encore quasiment d'une réciproque au sens propre du terme.

Les fonctions ψ et l'exposant μ_ψ seront davantage utilisés dans le paragraphe 12.

Mais ces deux exemples de réciproques de critères "à la Nesterenko" ne concernent qu'une seule variable : les formes linéaires construites sont des formes linéaires en 1 et ξ seulement. Or, nous avons besoin de formes en plusieurs variables $1, \xi_1, \ldots, \xi_{p-1}$. L'article [FHKL13] de Fischler, Hussain, Kristensen et Levesley contient une réciproque au critère de Nesterenko à plusieurs variables, valables presque partout au sens de la mesure de Lebesgue. On en énonce ici un cas particulier qui correspond au cadre et aux relations de cette thèse pour faciliter la comparaison.

Théorème 8.4 (Fischler-Hussain-Kristensen-Levesley, 2013) **:**
Soit $\tau_1, \dots, \tau_{p-1}, \beta_1, \dots, \beta_{p-1}, \varepsilon$ des réels tels que $\tau_1 > 0, \dots, \tau_{p-1} > 0, \varepsilon > 0$,

$$\tau_1 + \cdots + \tau_{p-1} \leq 1 \text{ et } \beta_1 + \cdots + \beta_{p-1} = (1+\varepsilon)(p-1)\,.$$

Alors, pour presque tout $(p-1)$-uplet $(\mathbf{e}_1, \dots, \mathbf{e}_{p-1}) \in (\mathbb{R}^p)^{p-1}$ au sens de la mesure de Lebesgue, pour tout Q assez grand, il existe des formes linéaires $L_Q^{(1)}, \dots, L_Q^{(p)} \in (\mathbb{R}^p)^*$ telles que :

1. Pour tout $j \in \{1, \dots, p\}$, $L_Q^{(j)}$ est à coefficients entiers ;
2. $L_Q^{(1)}, \dots, L_Q^{(p)}$ sont linéairement indépendantes ;
3. Pour tout $j \in \{1, \dots, p\}$, $\|L_Q^{(j)}\| \leq C_1 Q$;
4. Pour tout $i \in \{1, \dots, p-1\}$ et tout $j \in \{1, \dots, p\}$, on a

$$C_2 Q^{-\tau_i} (\log Q)^{\beta_i - (1+\varepsilon)p} \leq |L_Q^{(j)}(\mathbf{e}_i)| \leq C_3 Q^{-\tau_i} (\log Q)^{\beta_i},$$

où C_1, C_2 et C_3 sont des constantes strictement positives ne dépendant pas de Q.

Ce résultat fournit une sorte de réciproque (valable presque partout) aux propositions 6.1 et 5.1 des § 5 et 6 lorsque Λ_n est le réseau des formes linéaires à coefficients entiers (voir la première partie de la proposition 8.6 ci-dessous).

Toujours dans la direction de formes linéaires en plusieurs variables, Chantanasiri démontre également (§ 3 de [Cha12]) une sorte de réciproque sous une hypothèse très forte (une mesure d'indépendance linéaire). On donne ici une version de cette réciproque dans une formulation proche des énoncés de cette thèse et plus particulièrement du critère de Nesterenko originel 1.8 de l'introduction page 10, c'est-à-dire qu'on spécifie son résultat (proposition 3.3 de [Cha12]) au cas $Q_n = p\beta^n$ et $B = \beta^r U / c_0$:

Proposition 8.5 (Chatanasiri, 2012) :
Soient $\xi_0, \dots, \xi_{p-1}$ des réels linéairement indépendants sur $\mathbb{Q}$ tels qu'il existe $c_0 > 0$ tel que pour tout $\ell = (\ell_0, \dots, \ell_{p-1}) \in \mathbb{Z}^{p-1} \setminus \{\mathbf{0}\}$, on ait $|\ell_0\xi_0 + \cdots + \ell_{p-1}\xi_{p-1}| \geq c_0 \left(\max_{0\leq i\leq p-1} |\ell_i|\right)^{1-p}$. On pose $U = |\xi_0| + \cdots + |\xi_{p-1}|$ et $\beta > 1$.
Alors, pour tout $n \in \mathbb{N}$, il existe une forme linéaire non nulle à coefficients entiers en p variables

$$L_n(\mathbf{X}) = \ell_{0,n}X_0 + \cdots + \ell_{p-1,n}X_{p-1}$$

telle que

$$\sum_{i=0}^{p-1} |\ell_{i,n}| \leq p\beta^n, \quad 0 < |L_n(\boldsymbol{\xi})| \leq U\beta^{-n(p-1)}, \quad \frac{|L_{n-1}(\boldsymbol{\xi})|}{|L_n(\boldsymbol{\xi})|} \leq \frac{U\beta^r}{c_0}.$$

On adopte ici la même approche en essayant d'établir une sorte de réciproque au théorème 5.1. On montre que sous la condition de l'existence de p-uplets $\mathbf{a} = (a_1, \dots, a_p) \in \Lambda^{\perp}_{\Phi(Q)} \setminus \{\mathbf{0}\}$ avec $|a_i| \leq Q^{\tau_i - \varepsilon}$ et $|a_1\xi_1 + \cdots + a_{p-1}\xi_{p-1} + a_p| \leq Q^{-1-\varepsilon}$ pour un $\varepsilon > 0$ donné et une infinité de Q (ce qui est presque le contraire de la conclusion des théorèmes 5.1 et 6.1), on peut construire des entiers $\ell_{i,n}$ non tous nuls tels que $|\ell_{i,n} - \xi_i\ell_{p,n}| \leq Q_n^{-\tau_i+o(1)}$ et $|\ell_{p,n}| \leq Q_n^{1+o(1)}$ pour n assez grand (voir la proposition 8.7 ci-dessous).

On commence par montrer que les hypothèses du théorème 5.1 ne sont pas trop fortes et qu'il est possible de construire, pour des réels $\xi_1, \dots, \xi_{p-1}$ donnés, des formes linéaires satisfaisant à la plupart des conditions du théorème 5.1. Puis, on se pose la même question vis à vis des a_i. En effet, chacun des énoncés des parties précédentes affirme qu'une forme linéaire en $1, \xi_1, \dots, \xi_{p-1}$, à coefficients entiers pas trop gros et suffisamment divisibles, ne peut être trop petite. On démontre dans la proposition 8.6 ci-dessous que cette conclusion n'est pas triviale, c'est-à-dire que (pour certaines valeurs des paramètres) une telle forme linéaire peut être très petite.

Proposition 8.6 (Construction des $\ell_{i,n}$ et a_i) :
Soient $\xi_1, \ldots, \xi_{p-1}$ des réels quelconques et $\tau_1, \ldots, \tau_{p-1}$ des réels > -1.
On se donne pour tout $i \in \{1, \ldots, p\}$, et tout $n \in \mathbb{N}$, $\delta_{i,n} \in \mathbb{N}^*$ avec $\delta_{i,n} | \delta_{i,n+1}$. On note $\Lambda_n = \delta_{1,n}\mathbb{Z} \oplus \cdots \oplus \delta_{p,n}\mathbb{Z}$ le réseau qu'ils définissent. On considère aussi $(Q_n)_{n\in\mathbb{N}}$ une suite d'entiers strictement croissante avec $Q_{n+1} = Q_n^{1+o(1)}$. On suppose que pour tout $i \in \{1, \ldots, p\}$, il existe $\gamma_i \in \mathbb{R}$ tel que $\delta_{i,n} = Q_n^{\gamma_i + o(1)}$. Alors :

- Si
$$\gamma_p + \sum_{\substack{1\le i\le p-1\\ \tau_i+\gamma_i\ge 0}} \tau_i + \gamma_i \le 1\,,$$
alors il existe, à partir d'un certain rang, $(\ell_{1,n}, \ldots, \ell_{p,n}) \in \Lambda_n \setminus \{\mathbf{0}\}$ tel que $|\ell_{p,n}\xi_i - \ell_{i,n}| \le Q_n^{-\tau_i + o(1)}$ pour $i \in \{1, \ldots, p-1\}$ et $|\ell_{p,n}| \le Q_n^{1+o(1)}$.
- Si
$$\gamma_p + \sum_{\substack{1\le i\le p-1\\ \tau_i+\gamma_i\ge 0}} \tau_i + \gamma_i > 1\,, \tag{13}$$
alors, pour tout $\varepsilon > 0$ suffisamment petit et pour tout Q suffisamment grand par rapport à ε, on a $\Lambda^{\perp}_{\Phi(Q)} \cap \mathcal{C}'_{Q,\varepsilon} \neq \{\mathbf{0}\}$.

Démonstration. On note $J = \Big\{j \in \{1, \ldots, p-1\}, \tau_j + \gamma_j \ge 0\Big\}$ et $\alpha = p - |J| \in \{1, \ldots, p\}$. On a donc $|J| = p - \alpha$. On considère aussi le réseau $\Lambda'_n = \bigoplus_{j\in J\cup\{p\}} \delta_{j,n}\mathbb{Z} \subset \mathbb{R}^{J\cup\{p\}}$ qui est de déterminant $\delta_{p,n}\prod_{j\in J}\delta_{j,n} = Q_n^{\gamma_p + \sum_{j\in J}\gamma_j + o(1)}$.

Commençons par démontrer le premier point.

On considère $K_n = \Big\{\mathbf{x} \in \mathbb{R}^{J\cup\{p\}}, |x_p\xi_j - x_j| \le \delta_{j,n}Q_n^{-\tau_j-\gamma_j}, j \in J, \text{ et } |x_p| \le \delta_{p,n}Q_n^{1-\gamma_p}\Big\}$. K_n est un compact convexe, symétrique, centré en l'origine et de volume $2^{p-\alpha+1}\det(\Lambda'_n)Q_n^{1-\gamma_p-\sum_{j\in J}\tau_j+\gamma_j}$.

En utilisant la relation (13), le théorème de Minkowski assure l'existence de $(\ell_{j,n})_{j\in J\cup\{p\}} \in \Lambda'_n$ tel que $|\ell_{p,n}\xi_j - \ell_{j,n}| \le Q_n^{-\tau_j + o(1)}$ pour tout $j \in J$ et $|\ell_{p,n}| \le Q_n^{1+o(1)}$.

Soit $\varepsilon > 0$. Pour $j \in \{1, \ldots, p-1\} \setminus J$, on a $\tau_j + \gamma_j < 0$ donc $\delta_{j,n} + 1 < 2Q_n^{-\tau_j - \varepsilon}$ pour n assez grand. L'intervalle $[\ell_{p,n}\xi_j - Q_n^{-\tau_j-\varepsilon}, \ell_{p,n}\xi_j + Q_n^{-\tau_j-\varepsilon}]$ contient donc un multiple entier de $\delta_{j,n}$ que l'on choisit pour $\ell_{j,n}$. Cela termine la preuve du premier point.

Montrons maintenant le deuxième point de la proposition.

Soit $\varepsilon > 0$ tel que
$$\gamma_p + \sum_{\substack{1\le i\le p-1\\ \tau_i+\gamma_i\ge 0}} \tau_i + \gamma_i > 1 + (p - \alpha + 2)\varepsilon\,.$$

On considère le compact

$$K_Q = \left\{ \mathbf{a} \in \mathbb{R}^{J \cup \{p\}}, \forall j \in J, |a_j| \leq Q^{\tau_j - \varepsilon} \text{ et } \left| \sum_{j \in J} a_j \xi_j + a_p \right| \leq Q^{-1-\varepsilon} \right\}.$$

Alors K_Q est un compact convexe, symétrique, centré en l'origine, de volume $2^{p-\alpha+1} Q^{\sum_{j \in J} \tau_j - 1 - (p-\alpha+1)\varepsilon}$.

$\Lambda'^{\perp}_{\Phi(Q)}$, le réseau dual de $\Lambda'_{\Phi(Q)}$, est de déterminant inverse de celui de $\Lambda'_{\Phi(Q)}$, c'est-à-dire égal à $\prod_{j \in J \cup \{p\}} \frac{1}{\delta_{j,\Phi(Q)}} = Q_{\Phi(Q)}^{-\gamma_p - \sum_{j \in J} \gamma_j + o(1)} \leq Q^{-\gamma_p - \sum_{j \in J} \gamma_j + \varepsilon}$, si Q est assez grand.

On a alors :

$$Q^{\sum_{j \in J} \tau_j - 1 - (p-\alpha+1)\varepsilon} > Q^{-\gamma_p - \sum_{j \in J} \gamma_j + \varepsilon} > \det(\Lambda'^{\perp}_{\Phi(Q)}),$$

puisque $\gamma_p + \sum_{j \in J} (\tau_j + \gamma_j) > 1 + (p - \alpha + 2)\varepsilon$.

Donc $\mathrm{Vol}(K_Q) > 2^{p-\alpha+1} \det(\Lambda'^{\perp}_{\Phi(Q)})$ et le théorème de Minkowski assure l'existence d'un $(a_j)_{j \in J \cup \{p\}} \in K_Q \cap \Lambda'^{\perp}_{\Phi(Q)}$ non nul. On pose $a_j = 0$ pour $j \in \{1, \ldots, p-1\} \setminus J$, ce qui termine la démonstration. □

Le deuxième point de la proposition précédente permet donc de construire une forme linéaire en $1, \xi_1, \ldots, \xi_{p-1}$ qui contredit les conclusions des théorèmes 5.1 et 6.1 des parties 5 et 6. Ainsi, les conclusions de ces théorèmes impliquent que l'hypothèse (13) faite dans le second point de la proposition 8.6 est fausse. On peut alors appliquer le premier point et construire des $\ell_{i,n}$ qui vérifient une partie des hypothèses des théorèmes 5.1 et 6.1. Par conséquent, la combinaison des deux points de la proposition 8.6 ci-dessus permet d'établir une sorte de réciproque aux théorèmes 5.1 et 6.1.

Proposition 8.7 :
Soit $\xi_1, \ldots, \xi_{p-1}$ des réels quelconques et $\tau_1, \ldots, \tau_{p-1}$ des réels > -1. Soit $(Q_n)_{n \in \mathbb{N}}$ une suite d'entiers strictement croissante telle que $Q_{n+1} = Q_n^{1+o(1)}$. Soit p suites d'entiers $(\delta_{i,n})_{n \in \mathbb{N}}$ telles que pour tout $i \in \{1, \ldots, p\}$, tout $n \in \mathbb{N}$, $\delta_{i,n} | \delta_{i,n+1}$ et il existe $\gamma_i \in \mathbb{R}$ tel que $\delta_{i,n} = Q_n^{\gamma_i + o(1)}$. On pose, pour tout $n \in \mathbb{N}$, $\Lambda_n = \bigoplus_{i=1}^{p} \delta_{i,n} \mathbb{Z}$ le réseau qu'ils définissent.
Supposons qu'il existe $\varepsilon > 0$ et une infinité de Q tels que $\Lambda^{\perp}_{\Phi(Q)} \cap \mathcal{C}'_{Q,\varepsilon} = \{\mathbf{0}\}$.
Alors, pour tout n assez grand, il existe $\ell_{i,n} \in \delta_{i,n}\mathbb{Z}$ pour $\leq i \leq p$, non tous nuls, tels que, pour $\leq i \leq p-$, $|\ell_{i,n} - \ell_{p,n}\xi_i| \leq Q_n^{-\tau_i + o(1)}$ et $|\ell_{p,n}| \leq Q_n^{1+o(1)}$.

La double absence de relation de récurrence et d'indépendance linéaire entre des rangs successifs des formes linéaires construites, a tendance à éloigner cette réciproque du critère à la Siegel pour la rapprocher du critère à la Nesterenko. Toutefois dans le critère à la Nesterenko (théorème 5.1), les estimations sur les tailles des formes linéaires et de leurs coefficients sont exactes, alors que nous ne sommes en mesure ici d'obtenir que des majorations. Cette remarque a en revanche tendance à tirer cette réciproque vers le critère à la Siegel qui nécessite, lui, uniquement des majorations et non plus des estimations exactes.

Troisième partie

Application à $\zeta(2)$ et $\zeta(3)$

Dans la partie qui suit, on reprend (dans sa langue d'origine, à savoir l'anglais) l'article [DZ14] qui est en commun avec Zudilin et a été récemment soumis. La version soumise est disponible sur ma page web, et est également déposée sur ArXiV (arXiv :1401.5322) et HAL (HAL :hal-00933967). Toutefois quelques modifications ont été apportées. Dans la version donnée ici, ont été ajoutées de courtes démonstrations des lemmes de la partie 10.1 et 4 lemmes supplémentaires : les lemmes 12.5 et 12.6 dans le paragraphe 12.3 complètent et détaillent la preuve de la proposition 12.8 ; les deux autres à la fin du paragraphe 12.4 font une comparaison entre les majorations des exposants diophantiens considérés ici et les majorations obtenues par la preuve d'Apéry de l'irrationalité de $\zeta(2)$ et $\zeta(3)$. Les références à cette thèse dans l'article ont également été supprimées. On notera enfin dans la démonstration du théorème principal 9.1 dans la partie 11 l'ajout du lien avec les théorèmes énoncés dans les parties précédentes de cette thèse. Le paragraphe 13 page 76 a aussi été ajouté. On y détaille les démarches adoptées pour faire les calculs numériques nécessaires pour la démonstration du théorème 9.1 dans la section 11.

9 Introduction

It is known that the Riemann zeta function $\zeta(s)$ takes irrational values at positive even integers. This follows from Euler's evaluation $\zeta(s)/\pi^s \in \mathbb{Q}$ for $s = 2, 4, 6, \ldots$ and from the transcendence of π. Less is known about the values of $\zeta(s)$ at odd integers $s > 1$. Apéry was the first to establish the irrationality of such $\zeta(s)$: he proved [Apé79] in 1978 that $\zeta(3)$ is irrational. The next major step in the direction was made by Ball and Rivoal [BR01] in 2000: they showed that there are infinitely many odd integers at which Riemann zeta function is irrational. Shortly after, Rivoal demonstrates [Riv02] in 2001 that one of the nine numbers $\zeta(5), \zeta(7), \ldots, \zeta(21)$ is irrational, while the second author [Zud01] reduces the nine to four: he proves that at least one of the four numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$ and $\zeta(11)$ is irrational.

Already in 1978, Apéry constructs linear forms in 1 and $\zeta(2)$, as well as in 1 and $\zeta(3)$, with integer coefficients that produce the irrationality of the two zeta values in a quantitative form: the constructions imply upper bounds $\mu(\zeta(2)) < 11.850878\ldots$ and $\mu(\zeta(3)) < 13.41782\ldots$ for the irrationality measures. Recall that the irrationality exponent $\mu(\alpha)$ of a real irrational α is the supremum of the set of exponents μ for which the inequality $|\alpha - p/q| < q^{-\mu}$ has infinitely many solutions in rationals p/q. Hata improves the above mentioned results to $\mu(\zeta(2)) < 5.687$ in [Hat95, Addendum] and to $\mu(\zeta(3)) < 7.377956\ldots$ in [Hat00]. Then Rhin and Viola study a permutation group related to $\zeta(2)$ in [RV96] and show that $\mu(\zeta(2)) < 5.441243$. They later apply their new permutation group arithmetic method to $\zeta(3)$ as well and prove the upper bound $\mu(\zeta(3)) < 5.513891$. In an attempt to unify the achievements of Ball–Rivoal and of Rhin–Viola, the second author re-interpreted the constructions using the classical theory of hypergeometric functions and integrals [Zud04]. In his recent work [Zud14] (see also [Zud13]), he uses the permutation group arithmetic method and a hypergeometric construction, closely related to the one in this paper, to sharpen the earlier irrationality exponent of $\zeta(2)$ by proving that $\mu(\zeta(2)) \le 5.09541178\ldots$.

In this paper, we construct simultaneous rational approximations to both $\zeta(2)$ and $\zeta(3)$

using hypergeometric tools, and establish from them a lower bound for $\mathbb{Z}$-linear combinations of 1, $\zeta(2)$ and $\zeta(3)$ under some strong divisibility conditions on the coefficients. Namely, we prove the following result, stated as a theorem in the introduction of this thesis.

Theorem 9.1 :
Let η and ε be positive real numbers. For m sufficiently large with respect to ε and η, let $(a_0, a_1, a_2) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ be such that

(i) $D_m^2 D_{2m} a_0 \in \mathbb{Z}$, $D_m a_1 \in \mathbb{Z}$ and $\dfrac{D_{2m}}{D_m} a_2 \in \mathbb{Z}$, where D_m denotes the least common multiple of $1, 2, \dots, m$; and

(ii) $|a_0|, |a_1|, |a_2| \le e^{-(\tau_0+\varepsilon)m}$ hold with $\tau_0 = 0.899668635\dots$.

Then $|a_0 + a_1\zeta(2) + a_2\zeta(3)| > e^{-(s_0+\eta)m}$ with $s_0 = 6.770732145\dots$.

Theorem 9.1 contains the irrationality of both $\zeta(2)$ and $\zeta(3)$, because $\tau_0 < 1$. Namely, taking

$$a_0 = \frac{-p}{D_m}, \quad a_1 = \frac{q}{D_m} \quad \text{and} \quad a_2 = 0$$

shows that $\zeta(2) \ne p/q$, while the choice

$$a_0 = \frac{-D_m p}{D_{2m}}, \quad a_1 = 0 \quad \text{and} \quad a_2 = \frac{D_m q}{D_{2m}}$$

implies that $\zeta(3) \ne p/q$. The theorem does not give however the expected linear independence of 1, $\zeta(2)$ and $\zeta(3)$: it remains an open problem.

Our proof of Theorem 9.1 heavily rests upon a general version of hypergeometric construction of linear forms in 1 and $\zeta(2)$ on one hand, and in 1 and $\zeta(3)$ on the other hand; some particular instances of this construction were previously outlined in [Zud11]. More precisely, the linear forms $r_n = q_n\zeta(2) - p_n$ and $\hat{r}_n = \hat{q}_n\zeta(3) - \hat{p}_n$ we construct in the proof are hypergeometric-type series that depend on certain sets of auxiliary integer parameters. Permuting parameters in the sets allows us to gain p-adic information about the coefficients q_n, $\hat{q}_n$, p_n and $\hat{p}_n$. In addition, a classical transformation from the theory of hypergeometric functions implies that $q_n = \hat{q}_n$. The latter fact leads us to *simultaneous* rational approximations $r_n = q_n\zeta(2) - p_n$ and $\hat{r}_n = q_n\zeta(3) - \hat{p}_n$ to $\zeta(2)$ and $\zeta(3)$, with the following arithmetical and asymptotic properties:

$$\hat{\Phi}_n^{-1} q_n,\ \hat{\Phi}_n^{-1} D_{8n} D_{16n} p_n,\ \hat{\Phi}_n^{-1} D_{8n}^3 \hat{p}_n \in \mathbb{Z},$$

$$\lim_{n\to\infty} \frac{\log \hat{\Phi}_n}{n} = \varphi = 5.70169601\dots,$$

where $\hat{\Phi}_n$ is an explicit product over primes, and

$$\limsup_{n\to\infty} \frac{\log|r_n|}{n} = \limsup_{n\to\infty} \frac{\log|\hat{r}_n|}{n} = -\rho = -19.10095491\dots,$$

$$\lim_{n\to\infty} \frac{\log|q_n|}{n} = \kappa = 27.86755317\dots.$$

Finally, executing the Gosper–Zeilberger algorithm of creative telescoping we find out a recurrence relation satisfied by the linear forms r_n and $\hat{r}_n$. Together with a standard argument using the nonvanishing determinants formed from the coefficients of the forms, we then deduce Theorem 9.1 (some further computational details can be found in section 13). Note that

$$\tau_0 = \frac{1}{8}(32-\varphi-\rho) = 0.899668635\ldots \quad \text{and} \quad s_0 = \frac{1}{8}(32-\varphi+\kappa) = 6.770732145\ldots, \tag{14}$$

and the integer m from Theorem 9.1 is essentially $8n$.

In order to accommodate the atypical simultaneous approximations in Theorem 9.1 as well as to relate them to the context of previous results listed in the beginning of the section, we define a new diophantine exponent $s_\tau(\xi_1, \xi_2)$ of two real numbers ξ_1 and ξ_2, a characteristic of simultaneous irrationality of the numbers which depends on an additional parameter τ. With this notion in mind, we restate Theorem 9.1 as $s_{\tau_0}(\zeta(2), \zeta(3)) \le S$. Exploiting further the properties of the exponent, we demonstrate in Proposition 12.7 that if 1, $\zeta(2)$ and $\zeta(3)$ are linearly dependent aver $\mathbb{Q}$ (which seems very unlikely), then Theorem 9.1 holds with $s_0 = 6 - \tau_0 = 5.10033\ldots$ unless $\zeta(2)$ and $\zeta(3)$ belong to a certain set of Lebesgue measure 0.

In § 10 we introduce hypergeometric tools which depend on some parameters that lead to $\mathbb{Q}$-linear forms in 1 and $\zeta(2)$ on one hand, and in 1 and $\zeta(3)$ on the other, the forms having some common asymptotic properties.

In § 11 we specialise the parameters of the previous part to have the coefficients of $\zeta(2)$ and $\zeta(3)$ coincide. From this specialisation we derive the main theorem using recurrence relations satisfied by the linear forms and their coefficients.

In the final part, § 12, we introduce a new diophantine exponent. Some basic properties of this exponent are given, and it is compared to the irrationality exponents previously known. Then the main result is restated in terms of this diophantine exponent in Theorem 12.1, for consistency with previous results in the subject.

10 Hypergeometric series

In what follows, we always assume standard hypergeometric notation [Sla66]. For $n \in \mathbb{N}$, the Pochhammer symbol is given by

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = \prod_{k=0}^{n-1}(a+k),$$

with the convention $(a)_0 = 1$, while the generalized hypergeometric function is defined by the series

$${}_{p+1}F_p\left(\begin{matrix} a_0,\ a_1,\ \ldots,\ a_p \\ b_1,\ \ldots,\ b_p \end{matrix}\ \middle|\ z\right) = \sum_{n=0}^{\infty} \frac{(a_0)_n (a_1)_n \cdots (a_p)_n}{n!\,(b_1)_n \cdots (b_p)_n}\, z^n.$$

10.1 Integer-valued polynomials

This section discusses results about integer-valued polynomials.

Lemma 10.1 :
For $\ell = 0, 1, 2, \dots,$

$$\frac{1}{2\pi i}\int_{1/2-i\infty}^{1/2+i\infty}\left(\frac{\pi}{\sin \pi t}\right)^2 \frac{(t-1)(t-2)\cdots(t-\ell)}{\ell!}\,\mathrm{d}t = \frac{(-1)^\ell}{\ell+1}. \tag{15}$$

Proof. The integrand is

$$\frac{1}{\ell!}\left(\frac{\pi}{\sin \pi t}\right)^2 \frac{\Gamma(t)}{\Gamma(t-\ell)} = \frac{(-1)^\ell}{\ell!}\,\Gamma(t)^2\Gamma(1-t)\,\Gamma(1+\ell-t);$$

the evaluation in (15) follows from Barnes's first lemma [Sla66, Section 4.2.1]. □

Lemma 10.2 :
Given $b < a$ integers, set

$$R(t) = R(a, b; t) = \frac{(t+b)(t+b+1)\cdots(t+a-1)}{(a-b)!}.$$

Then

$$R(k) \in \mathbb{Z}, \quad D_{a-b}\cdot\frac{\mathrm{d}R(t)}{\mathrm{d}t}\bigg|_{t=k} \in \mathbb{Z} \quad \text{and} \quad D_{a-b}\cdot\frac{R(k)-R(\ell)}{k-\ell} \in \mathbb{Z}$$

for any $k, \ell \in \mathbb{Z}$, $\ell \neq k$.

Proof. Denote by $m = b - a$ the degree of the polynomial $R(t)$. The first two families of inclusions are classical (see for instance [Zud04]). For the remaining one, introduce the polynomial $P(t) = (R(t) - R(\ell))/(t-\ell)$ of degree $m-1$. As $D_m \cdot 1/(k-\ell)$ is an integer for $k = \ell+1, \ell+2, \dots, \ell+m$ as well as $R(k) - R(\ell) \in \mathbb{Z}$, we deduce that $D_m \cdot P(k) \in \mathbb{Z}$ for those values of k. This means that the polynomial $D_m P(t)$ of degree $m-1$ assumes integer values at m successive integers; by [PS98, Division 8, Problem 87] this polynomial is integer-valued. □

Lemma 10.3 :
Let $R(t)$ be a product of several integer-valued polynomials

$$R_j(t) = R(a_j, b_j; t) = \frac{(t+b_j)(t+b_j+1)\cdots(t+a_j-1)}{(a_j-b_j)!}, \quad \text{where } b_j < a_j,$$

and $m = \max_j\{a_j - b_j\}$. Then

$$R(k) \in \mathbb{Z}, \quad D_m \cdot \frac{\mathrm{d}R(t)}{\mathrm{d}t}\bigg|_{t=k} \in \mathbb{Z} \quad \text{and} \quad D_m \cdot \frac{R(k)-R(\ell)}{k-\ell} \in \mathbb{Z} \tag{16}$$

for any $k, \ell \in \mathbb{Z}$, $\ell \neq k$.

Proof. It is sufficient to establish the result for a product of just two polynomials $R(t)$ and $\widetilde{R}(t)$ satisfying the inclusions (16), and then use mathematical induction on the number of such factors. We have

$$\begin{aligned}
\big(R(t)\widetilde{R}(t)\big)\big|_{t=k} &= R(k)\widetilde{R}(k) \in \mathbb{Z}, \\
D_m \frac{\mathrm{d}(R(t)\widetilde{R}(t))}{\mathrm{d}t}\bigg|_{t=k} &= R(k) \cdot D_m \frac{\mathrm{d}\widetilde{R}(t)}{\mathrm{d}t}\bigg|_{t=k} + D_m \frac{\mathrm{d}R(t)}{\mathrm{d}t}\bigg|_{t=k} \cdot \widetilde{R}(k) \in \mathbb{Z} \\
D_m \frac{R(k)\widetilde{R}(k) - R(\ell)\widetilde{R}(\ell)}{k-\ell} &= D_m \frac{R(k)-R(\ell)}{k-\ell} \cdot \widetilde{R}(k) + R(\ell) \cdot D_m \frac{\widetilde{R}(k)-\widetilde{R}(\ell)}{k-\ell} \in \mathbb{Z},
\end{aligned}$$

and the result follows. □

10.2 Construction of linear forms in 1 and $\zeta(2)$

The construction in this subsection is a general case of the one considered in [Zud07, Section 2].

For a set of parameters

$$(a, b) = \begin{pmatrix} a_1, & a_2, & a_3, & a_4 \\ b_1, & b_2, & b_3, & b_4 \end{pmatrix}$$

subject to the conditions

$$\begin{gathered}
b_1, b_2, b_3 \leq a_1, a_2, a_3, a_4 < b_4, \\
d = (a_1 + a_2 + a_3 + a_4) - (b_1 + b_2 + b_3 + b_4) \geq 0,
\end{gathered} \tag{17}$$

define the rational function

$$\begin{aligned}
R(t) = R(a, b; t) &= \frac{(t+b_1)\cdots(t+a_1-1)}{(a_1-b_1)!} \cdot \frac{(t+b_2)\cdots(t+a_2-1)}{(a_2-b_2)!} \\
&\quad \times \frac{(t+b_3)\cdots(t+a_3-1)}{(a_3-b_3)!} \cdot \frac{(b_4-a_4-1)!}{(t+a_4)\cdots(t+b_4-1)} \qquad (18) \\
&= \Pi(a, b) \cdot \frac{\Gamma(t+a_1)\,\Gamma(t+a_2)\,\Gamma(t+a_3)\,\Gamma(t+a_4)}{\Gamma(t+b_1)\,\Gamma(t+b_2)\,\Gamma(t+b_3)\,\Gamma(t+b_4)}, \qquad (19)
\end{aligned}$$

where
$$\Pi(a,b) = \frac{(b_4 - a_4 - 1)!}{(a_1 - b_1)!\,(a_2 - b_2)!\,(a_3 - b_3)!}.$$
We also introduce the ordered versions $a_1^* \le a_2^* \le a_3^* \le a_4^*$ of the parameters a_1, a_2, a_3, a_4 and $b_1^* \le b_2^* \le b_3^*$ of b_1, b_2, b_3, so that $\{a_1^*, a_2^*, a_3^*, a_4^*\}$ coincides with $\{a_1, a_2, a_3, a_4\}$ and $\{b_1^*, b_2^*, b_3^*\}$ coincides with $\{b_1, b_2, b_3\}$ as multi-sets (that is, sets with possible repetition of elements). Then $R(t)$ has poles at $t = -k$ where $k = a_4^*, a_4^*+1, \dots, b_4 - 1$, zeroes at $t = -\ell$ where $\ell = b_1^*, b_1^*+1, \dots, a_3^*-1$, and double zeroes at $t = -\ell$ where $\ell = b_2^*, b_2^*+1, \dots, a_2^*-1$.

Decomposing $R(t)$ into the sum of partial fractions, we get
$$R(t) = \sum_{k=a_4^*}^{b_4-1} \frac{C_k}{t+k} + P(t), \tag{20}$$
where $P(t)$ is a polynomial of which the degree d is defined in (17) and
$$\begin{aligned} C_k &= \big(R(t)(t+k)\big)|_{t=-k} \\ &= (-1)^{d+b_4+k} \binom{k-b_1}{k-a_1}\binom{k-b_2}{k-a_2}\binom{k-b_3}{k-a_3}\binom{b_4-a_4-1}{k-a_4} \in \mathbb{Z} \end{aligned} \tag{21}$$
for $k = a_4^*, a_4^*+1, \dots, b_4 - 1$.

Lemma 10.4 :
Set $c = \max\{a_1 - b_1, a_2 - b_2, a_3 - b_3\}$. Then $D_c P(t)$ is an integer-valued polynomial of degree d.

Proof. Write $R(t) = R_1(t)R_2(t)$, where
$$R_1(t) = \frac{\prod_{j=b_1}^{a_1-1}(t+j)}{(a_1-b_1)!} \cdot \frac{\prod_{j=b_2}^{a_2-1}(t+j)}{(a_2-b_2)!} \cdot \frac{\prod_{j=b_3}^{a_3-1}(t+j)}{(a_3-b_3)!}$$
is the product of three integer-valued polynomials and
$$R_2(t) = \frac{(b_4-a_4-1)!}{\prod_{j=a_4}^{b_4-1}(t+j)} = \sum_{k=a_4}^{b_4-1} \frac{(-1)^{k-a_4}\binom{b_4-a_4-1}{k-a_4}}{t+k}.$$

It follows from Lemma 10.3 that
$$\begin{aligned} &D_c \cdot \frac{\mathrm{d}R_1(t)}{\mathrm{d}t}\bigg|_{t=j} \in \mathbb{Z} \quad \text{for } j \in \mathbb{Z} \quad \text{and} \\ &D_c \cdot \frac{R_1(j) - R_1(m)}{j-m} \in \mathbb{Z} \quad \text{for } j, m \in \mathbb{Z},\ j \ne m. \end{aligned} \tag{22}$$

Furthermore, note that
$$\begin{aligned} C_k &= R_1(-k) \cdot \big(R_2(t)(t+k)\big)\big|_{t=-k} \\ &= R_1(-k) \cdot (-1)^{k-a_4} \binom{b_4-a_4-1}{k-a_4} \quad \text{for } k \in \mathbb{Z}, \end{aligned}$$

and the expression in fact vanishes if k is outside the range $a_4^* \le k \le b_4 - 1$.

For $\ell \in \mathbb{Z}$ we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big(R(t)(t+\ell)\big)\Big|_{t=-\ell} &= \frac{\mathrm{d}}{\mathrm{d}t}\big(R_1(t)\cdot R_2(t)(t+\ell)\big)\Big|_{t=-\ell} \\
&= \frac{\mathrm{d}R_1(t)}{\mathrm{d}t}\Big|_{t=-\ell}\cdot\big(R_2(t)(t+\ell)\big)\big|_{t=-\ell} + R_1(-\ell)\cdot\frac{\mathrm{d}}{\mathrm{d}t}\big(R_2(t)(t+\ell)\big)\Big|_{t=-\ell} \\
&= \frac{\mathrm{d}R_1(t)}{\mathrm{d}t}\Big|_{t=-\ell}\cdot(-1)^{\ell-a_4}\binom{b_4-a_4-1}{\ell-a_4} \\
&\qquad + R_1(-\ell)\cdot\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=a_4}^{b_4-1}(-1)^{k-a_4}\binom{b_4-a_4-1}{k-a_4}\left(1-\frac{-\ell+k}{t+k}\right)\Big|_{t=-\ell} \\
&= \frac{\mathrm{d}R_1(t)}{\mathrm{d}t}\Big|_{t=-\ell}\cdot(-1)^{\ell-a_4}\binom{b_4-a_4-1}{\ell-a_4} + R_1(-\ell)\sum_{\substack{k=a_4\\k\ne\ell}}^{b_4-1}\frac{(-1)^{k-a_4}\binom{b_4-a_4-1}{k-a_4}}{-\ell+k}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=a_4^*}^{b_4-1}\frac{C_k}{t+k}\cdot(t+\ell)\right)\Big|_{t=-\ell} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=a_4}^{b_4-1}\frac{C_k}{t+k}\cdot(t+\ell)\right)\Big|_{t=-\ell} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=a_4}^{b_4-1}C_k\left(1-\frac{-\ell+k}{t+k}\right)\Big|_{t=-\ell} = \sum_{\substack{k=a_4\\k\ne\ell}}^{b_4-1}\frac{C_k}{-\ell+k} \\
&= \sum_{\substack{k=a_4\\k\ne\ell}}^{b_4-1}\frac{R_1(-k)\cdot(-1)^{k-a_4}\binom{b_4-a_4-1}{k-a_4}}{-\ell+k}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
P(-\ell) = \frac{\mathrm{d}}{\mathrm{d}t}\big(P(t)(t+\ell)\big)\big|_{t=-\ell} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(R(t)(t+\ell) - \sum_{k=a_4^*}^{b_4-1}\frac{C_k}{t+k}\cdot(t+\ell)\right)\Big|_{t=-\ell} \\
&= \frac{\mathrm{d}R_1(t)}{\mathrm{d}t}\Big|_{t=-\ell}\cdot(-1)^{\ell-a_4}\binom{b_4-a_4-1}{\ell-a_4} \\
&\qquad + \sum_{\substack{k=a_4\\k\ne\ell}}^{b_4-1}(-1)^{k-a_4}\binom{b_4-a_4-1}{k-a_4}\frac{R_1(-\ell)-R_1(-k)}{-\ell+k},
\end{aligned} \tag{23}$$

and this implies, on the basis of the inclusions (22) above, that $D_c P(-\ell) \in \mathbb{Z}$ for all $\ell \in \mathbb{Z}$. $\square$

Finally, define the quantity

$$r(a,b) = \frac{(-1)^d}{2\pi i}\int_{C-i\infty}^{C+i\infty}\left(\frac{\pi}{\sin\pi t}\right)^2 R(a,b;t)\,\mathrm{d}t, \tag{24}$$

where C is arbitrary from the interval $-a_2^* < C < 1 - b_2^*$. The definition does not depend on the choice of C, as the integrand does not have singularities in the strip $-a_2^* < \operatorname{Re} t < 1 - b_2^*$.

Proposition 10.5 :
We have

$$r(a,b) = q(a,b)\zeta(2) - p(a,b), \qquad \text{with} \quad q(a,b) \in \mathbb{Z}, \quad D_{c_1} D_{c_2} p(a,b) \in \mathbb{Z}, \tag{25}$$

where

$$c_1 = \max\{a_1 - b_1, a_2 - b_2, a_3 - b_3, b_4 - a_2^* - 1\} \quad \text{and} \quad c_2 = \max\{d+1, b_4 - a_2^* - 1\}.$$

In addition,

$$\begin{aligned} q(a,b) \;&=\; (-1)^{b_4 - a_4^* - 1} \binom{a_4^* - b_1}{a_4^* - a_1} \binom{a_4^* - b_2}{a_4^* - a_2} \binom{a_4^* - b_3}{a_4^* - a_3} \binom{b_4 - a_4 - 1}{a_4^* - a_4} \\ &\quad \times {}_4F_3\left(\left. \begin{matrix} -(b_4 - a_4^* - 1),\ a_4^* - b_1 + 1,\ a_4^* - b_2 + 1,\ a_4^* - b_3 + 1 \\ a_4^* - a_1^* + 1,\ a_4^* - a_2^* + 1,\ a_4^* - a_3^* + 1 \end{matrix} \right| 1 \right), \end{aligned} \tag{26}$$

and the quantity $r(a,b)/\Pi(a,b)$ is invariant under any permutation of the parameters a_1, a_2, a_3, a_4.

Proof. We choose $C = 1/2 - a_2^*$ in (24) and write (20) as

$$R(t) = \sum_{k=a_4^*}^{b_4-1} \frac{C_k}{t+k} + \sum_{\ell=0}^{d} A_\ell P_\ell(t + a_2^*),$$

where

$$P_\ell(t) = \frac{(t-1)(t-2)\cdots(t-\ell)}{\ell!}$$

and $D_c A_\ell \in \mathbb{Z}$ in accordance with Lemma 10.4. Applying Lemma 10.1 we obtain

$$\begin{aligned} r(a,b) &= \frac{(-1)^d}{2\pi i} \int_{1/2 - i\infty}^{1/2 + i\infty} \left(\frac{\pi}{\sin \pi t} \right)^2 R(t - a_2^*)\, \mathrm{d}t \\ &= (-1)^d \sum_{m=1-a_2^*}^{\infty} \sum_{k=a_4^*}^{b_4-1} \frac{C_k}{(m+k)^2} + \sum_{\ell=0}^{d} \frac{(-1)^{d+\ell} A_\ell}{\ell+1} \\ &= \zeta(2) \cdot (-1)^d \sum_{k=a_4^*}^{b_4-1} C_k - (-1)^d \sum_{k=a_4^*}^{b_4-1} C_k \sum_{\ell=1}^{k-a_2^*} \frac{1}{\ell^2} + \sum_{\ell=0}^{d} \frac{(-1)^{d+\ell} A_\ell}{\ell+1}. \end{aligned} \tag{27}$$

This representation clearly implies that $r(a,b)$ has the desired form (25), while the hypergeometric form (26) follows from

$$q(a,b) = (-1)^d \sum_{k=a_4^*}^{b_4-1} C_k$$

and the explicit formula (21) for C_k. Finally, the invariance of $r(a,b)/\Pi(a,b)$ under permutations of a_1, a_2, a_3, a_4 follows from (19) and definition (24) of $r(a,b)$. $\square$

Assume that the parameters (a,b) are chosen in the following way:

$$\begin{aligned} &a_1 = \alpha_1 n + 1, \quad a_2 = \alpha_2 n + 1, \quad a_3 = \alpha_3 n + 1, \quad a_4 = \alpha_4 n + 1, \\ &b_1 = \beta_1 n + 1, \quad b_2 = \beta_2 n + 1, \quad b_3 = \beta_3 n + 1, \quad b_4 = \beta_4 n + 2, \end{aligned} \tag{28}$$

where the *fixed* integers α_j and β_j, $j = 1, \dots, 4$, satisfy

$$\begin{gathered} \beta_1, \beta_2, \beta_3 < \alpha_1, \alpha_2, \alpha_3, \alpha_4 < \beta_4, \\ \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 > \beta_1 + \beta_2 + \beta_3 + \beta_4. \end{gathered}$$

The quantities (25) in these settings become dependent on a single parameter $n = 0, 1, 2, \dots$, so we let $r_n = r(a,b)$, $q_n = q(a,b)$, $p_n = p(a,b)$ and identify the characteristics $c_1 = \gamma_1 n$ and $c_2 = \gamma_2 n$ of Proposition 10.5, where γ_1 and γ_2 are completely determined by α_j and β_j, $j = 1, \dots, 4$. The statement below is proven by standard techniques and is very similar to [Zud04, Lemmas 10–12].

Proposition 10.6 :
In the above notation, let τ_0, $\overline{\tau_0} \in \mathbb{C} \setminus \mathbb{R}$ and $\tau_1 \in \mathbb{R}$ be the zeroes of the cubic polynomial $\prod_{j=1}^{4}(\tau - \alpha_j) - \prod_{j=1}^{4}(\tau - \beta_j)$. Define

$$\begin{aligned} f_0(\tau) \qquad &= \textstyle\sum_{j=1}^{4}\big(\alpha_j \log(\tau - \alpha_j) - \beta_j \log(\tau - \beta_j)\big) \\ &\quad - \textstyle\sum_{j=1}^{3}(\alpha_j - \beta_j)\log(\alpha_j - \beta_j) + (\beta_4 - \alpha_4)\log(\beta_4 - \alpha_4). \end{aligned}$$

Then
$$\limsup_{n\to\infty} \frac{\log|r_n|}{n} = \operatorname{Re} f_0(\tau_0) \quad \text{and} \quad \lim_{n\to\infty} \frac{\log|q_n|}{n} = \operatorname{Re} f_0(\tau_1).$$
Furthermore,
$$\Phi_n^{-1} q_n,\ \Phi_n^{-1} D_{\gamma_1 n} D_{\gamma_2 n} p_n \in \mathbb{Z}$$
with
$$\Phi_n = \prod_{\substack{p \text{ prime} \\ p \le \min\{\gamma_1, \gamma_2\} n}} p^{\varphi(n/p)},$$
where
$$\begin{aligned} \varphi(x) = \max_{\alpha' = \sigma\alpha : \sigma \in \mathfrak{S}_4} \Big(&\lfloor(\beta_4 - \alpha_4)x\rfloor - \lfloor(\beta_4 - \alpha_4')x\rfloor \\ &- \textstyle\sum_{j=1}^{3}\big(\lfloor(\alpha_j - \beta_j)x\rfloor - \lfloor(\alpha_j' - \beta_j)x\rfloor\big)\Big), \end{aligned}$$
so that the maximum is taken over all permutations $(\alpha_1', \alpha_2', \alpha_3', \alpha_4')$ of $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, and we have
$$\lim_{n\to\infty} \frac{\log \Phi_n}{n} = \int_0^1 \varphi(x)\,\mathrm{d}\psi(x) - \int_0^{1/\min\{\gamma_1,\gamma_2\}} \varphi(x)\,\frac{\mathrm{d}x}{x^2},$$
where $\psi(x)$ is the logarithmic derivative of the gamma function.

Here and in what follows, the notation $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ is used for the floor and ceiling integer-part functions.

10.3 Construction of linear forms in 1 and $\zeta(3)$

The construction in this subsection depends on another set of integral parameters

$$(\hat{a}, \hat{b}) = \begin{pmatrix} \hat{a}_0,\ \hat{a}_1,\ \hat{a}_2,\ \hat{a}_3 \\ \hat{b}_0,\ \hat{b}_1,\ \hat{b}_2,\ \hat{b}_3 \end{pmatrix}$$

which satisfies the conditions

$$\begin{gathered} \tfrac{1}{2}\hat{b}_0,\ \hat{b}_1 \le \tfrac{1}{2}\hat{a}_0,\ \hat{a}_1,\ \hat{a}_2,\ \hat{a}_3 < \hat{b}_2,\ \hat{b}_3, \\ \hat{a}_0 + \hat{a}_1 + \hat{a}_2 + \hat{a}_3 \le \hat{b}_0 + \hat{b}_1 + \hat{b}_2 + \hat{b}_3 - 2. \end{gathered} \tag{29}$$

To this set we assign the rational function

$$\begin{aligned}\hat{R}(t)=\hat{R}(\hat{a},\hat{b};t)&=\frac{(2t+\hat{b}_0)(2t+\hat{b}_0+1)\cdots(2t+\hat{a}_0-1)}{(\hat{a}_0-\hat{b}_0)!}\cdot\frac{(t+\hat{b}_1)\cdots(t+\hat{a}_1-1)}{(\hat{a}_1-\hat{b}_1)!}\\&\qquad\times\frac{(\hat{b}_2-\hat{a}_2-1)!}{(t+\hat{a}_2)\cdots(t+\hat{b}_2-1)}\cdot\frac{(\hat{b}_3-\hat{a}_3-1)!}{(t+\hat{a}_3)\cdots(t+\hat{b}_3-1)} &(30)\\&=\hat{\Pi}(\hat{a},\hat{b})\cdot\frac{\Gamma(2t+\hat{a}_0)\,\Gamma(t+\hat{a}_1)\,\Gamma(t+\hat{a}_2)\,\Gamma(t+\hat{a}_3)}{\Gamma(2t+\hat{b}_0)\,\Gamma(t+\hat{b}_1)\,\Gamma(t+\hat{b}_2)\,\Gamma(t+\hat{b}_3)}, &(31)\end{aligned}$$

where

$$\hat{\Pi}(\hat{a},\hat{b})=\frac{(\hat{b}_2-\hat{a}_2-1)!\,(\hat{b}_3-\hat{a}_3-1)!}{(\hat{a}_0-\hat{b}_0)!\,(\hat{a}_1-\hat{b}_1)!}.$$

As in Section 10.2 we introduce the ordered versions $\hat{a}_1^*\le\hat{a}_2^*\le\hat{a}_3^*$ of the parameters $\hat{a}_1,\hat{a}_2,\hat{a}_3$ and $\hat{b}_2^*\le\hat{b}_3^*$ of $\hat{b}_2,\hat{b}_3$. Then this ordering and conditions (29) imply that $\hat{R}(t)=O(1/t^2)$ as $t\to\infty$, the rational function has poles at $t=-k$ for $\hat{a}_2^*\le k\le\hat{b}_3^*-1$, double poles at $t=-k$ for $\hat{a}_3^*\le k\le\hat{b}_2^*-1$, and double zeroes at $t=-\ell$ for $\max\{\lceil\hat{b}_0/2\rceil,\hat{b}_1\}\le\ell\le\min\{\lfloor(\hat{a}_0-1)/2\rfloor,\hat{a}_1^*-1\}$.

The partial-fraction decomposition of $\hat{R}(t)$ assumes the form

$$\hat{R}(t)=\sum_{k=\hat{a}_3^*}^{\hat{b}_2^*-1}\frac{A_k}{(t+k)^2}+\sum_{k=\hat{a}_2^*}^{\hat{b}_3^*-1}\frac{B_k}{t+k}, \tag{32}$$

where

$$\begin{aligned}A_k&=\big(\hat{R}(t)(t+k)^2\big)|_{t=-k}\\&=(-1)^{\hat{d}}\binom{2k-\hat{b}_0}{2k-\hat{a}_0}\binom{k-\hat{b}_1}{k-\hat{a}_1}\binom{\hat{b}_2-\hat{a}_2-1}{k-\hat{a}_2}\binom{\hat{b}_3-\hat{a}_3-1}{k-\hat{a}_3}\in\mathbb{Z} &(33)\end{aligned}$$

with $\hat{d}=\hat{a}_0+\hat{a}_1+\hat{a}_2+\hat{a}_3-\hat{b}_0-\hat{b}_1$, for $k=\hat{a}_3^*,\hat{a}_3^*+1,\dots,\hat{b}_2^*-1$ and, similarly,

$$B_k=\frac{\mathrm{d}}{\mathrm{d}t}\big(\hat{R}(t)(t+k)^2\big)|_{t=-k} \tag{34}$$

for $k=\hat{a}_2^*,\hat{a}_2^*+1,\dots,\hat{b}_3^*-1$. The inclusions

$$D_{\max\{\hat{a}_0-\hat{b}_0,\hat{a}_1-\hat{b}_1,\hat{b}_3^*-\hat{a}_2-1,\hat{b}_3^*-\hat{a}_3-1\}}\cdot B_k\in\mathbb{Z} \tag{35}$$

follow then from standard consideration; see, for example, Lemma 3 and the proof of Lemma 4 in [Zud04]. In addition,

$$\sum_{k=\hat{a}_2^*}^{\hat{b}_3^*-1}B_k=-\operatorname*{Res}_{t=\infty}\hat{R}(t)=0 \tag{36}$$

by the residue sum theorem.

The quantity of our interest in this section is

$$\hat r(\hat a,\hat b)=\frac{(-1)^{\hat d}}{4\pi i}\int_{C-i\infty}^{C+i\infty}\left(\frac{\pi}{\sin\pi t}\right)^2\hat R(\hat a,\hat b;t)\,\mathrm{d}t, \tag{37}$$

where C is arbitrary from the interval $-\min\{\hat a_0/2,\hat a_1^*\}<C<1-\max\{\hat b_0/2,\hat b_1\}$.

Proposition 10.7 :
We have

$$\hat r(\hat a,\hat b)=\hat q(\hat a,\hat b)\zeta(3)-\hat p(\hat a,\hat b),\qquad \text{with}\quad \hat q(\hat a,\hat b)\in\mathbb{Z},\quad 2D_{\hat c_1}D_{\hat c_2}^2\hat p(\hat a,\hat b)\in\mathbb{Z}, \tag{38}$$

where

$$\hat c_1=\max\{\hat a_0-\hat b_0,\,\hat a_1-\hat b_1,\,\hat b_3^*-\hat a_2-1,\,\hat b_3^*-\hat a_3-1,\,\hat b_2^*-\lceil\hat a_0/2\rceil-1,\,\hat b_2^*-\hat a_1^*-1\},$$
$$\hat c_2=\max\{\hat b_3^*-\lceil\hat a_0/2\rceil-1,\,\hat b_3^*-\hat a_1^*-1\}.$$

Furthermore,

$$\begin{aligned}\hat q(\hat a,\hat b)&=\tbinom{2\hat a_3^*-\hat b_0}{2\hat a_3^*-\hat a_0}\tbinom{\hat a_3^*-\hat b_1}{\hat a_3^*-\hat a_1}\tbinom{\hat b_2-\hat a_2-1}{\hat a_3^*-\hat a_2}\tbinom{\hat b_3-\hat a_3-1}{\hat a_3^*-\hat a_3}\\ &\times{}_5F_4\left(\left.\begin{matrix}-(\hat b_2-\hat a_3^*-1),\ -(\hat b_3-\hat a_3^*-1),\ \hat a_3^*-\hat b_1+1,\ \hat a_3^*-\frac12\hat b_0+\frac12,\ \hat a_3^*-\frac12\hat b_0+1\\ \hat a_3^*-\hat a_1^*+1,\ \hat a_3^*-\hat a_2^*+1,\ \hat a_3^*-\frac12\hat a_0+\frac12,\ \hat a_3^*-\frac12\hat a_0+1\end{matrix}\right|\,1\right),\end{aligned} \tag{39}$$

and the quantity $\hat r(\hat a,\hat b)/\hat\Pi(\hat a,\hat b)$ is invariant under any permutation of the parameters $\hat a_1,\hat a_2,\hat a_3$.

Proof. Denote $\hat a^*=\min\{\lceil\hat a_0/2\rceil,\hat a_1^*\}$ and choose $C=1/2-\hat a^*$ in (37) to write

$$\begin{aligned}\hat r(\hat a,\hat b)&=-\frac{(-1)^{\hat d}}{2}\sum_{m=1-\hat a^*}^{\infty}\left.\frac{\mathrm{d}\hat R(t)}{\mathrm{d}t}\right|_{t=m}\\ &=(-1)^{\hat d}\sum_{m=1-\hat a^*}^{\infty}\sum_{k=\hat a_3^*}^{\hat b_2^*-1}\frac{A_k}{(m+k)^3}+\frac{(-1)^{\hat d}}{2}\sum_{m=1-\hat a^*}^{\infty}\sum_{k=\hat a_2^*}^{\hat b_3^*-1}\frac{B_k}{(m+k)^2}\\ &=\zeta(3)\cdot(-1)^{\hat d}\sum_{k=\hat a_3^*}^{\hat b_2^*-1}A_k\\ &\qquad-(-1)^{\hat d}\sum_{k=\hat a_3^*}^{\hat b_2^*-1}A_k\sum_{\ell=1}^{k-\hat a^*}\frac{1}{\ell^3}-\frac{(-1)^{\hat d}}{2}\sum_{k=\hat a_2^*}^{\hat b_3^*-1}B_k\sum_{\ell=1}^{k-\hat a^*}\frac{1}{\ell^2},\end{aligned} \tag{40}$$

where equality (36) was used. In view of the inclusions (33), (35) the found representation of

$\hat{r}(\hat{a}, \hat{b})$ implies the form (38). The hypergeometric form (39) follows from

$$\hat{q}(\hat{a}, \hat{b}) = (-1)^{\hat{d}} \sum_{k=\hat{a}_3^*}^{\hat{b}_2^*-1} A_k$$

and the explicit formula (33) for A_k. Finally, the invariance of $\hat{r}(\hat{a}, \hat{b})/\hat{\Pi}(\hat{a}, \hat{b})$ under permutations of $\hat{a}_1, \hat{a}_2, \hat{a}_3$ follows from (31) and definition (37) of $\hat{r}(\hat{a}, \hat{b})$. $\square$

Similar to our choice in Section 10.2, we take the parameters $(\hat{a}, \hat{b})$ as follows:

$$\begin{aligned} &\hat{a}_0 = \hat{\alpha}_0 n + 2, \quad \hat{a}_1 = \hat{\alpha}_1 n + 1, \quad \hat{a}_2 = \hat{\alpha}_2 n + 1, \quad \hat{a}_3 = \hat{\alpha}_3 n + 1, \\ &\hat{b}_0 = \hat{\beta}_0 n + 2, \quad \hat{b}_1 = \hat{\beta}_1 n + 1, \quad \hat{b}_2 = \hat{\beta}_2 n + 2, \quad \hat{b}_3 = \hat{\beta}_3 n + 2, \end{aligned} \tag{41}$$

where the fixed integers $\hat{\alpha}_j$ and $\hat{\beta}_j$, $j = 0, \dots, 3$, satisfy

$$\begin{gathered} \tfrac{1}{2}\hat{\beta}_0, \hat{\beta}_1 < \tfrac{1}{2}\hat{\alpha}_0 \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3 < \hat{\beta}_2, \hat{\beta}_3, \\ \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = \hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3; \end{gathered}$$

note that the equality is assumed in the latter relation (compare to (29)) to simplify the asymptotic consideration in Proposition 10.8. The quantities (38) then depend on $n = 0, 1, 2, \dots$; we write $\hat{r}_n = \hat{r}(\hat{a}, \hat{b})$, $\hat{q}_n = \hat{q}(\hat{a}, \hat{b})$, $\hat{p}_n = \hat{p}(\hat{a}, \hat{b})$ and identify the characteristics $\hat{c}_1 = \hat{\gamma}_1 n$ and $\hat{c}_2 = \hat{\gamma}_2 n$ of Proposition 10.7. Proving the analytical part of the following statement is again similar to what is done in [Zud04, Lemma 12 or Lemma 20], while the arithmetic part follows from the results in [Zud04, Section 7] (cf. [Zud04, Lemma 19]).

Proposition 10.8 :
In the above notation, let $\hat{\tau}_0, \overline{\hat{\tau}_0} \in \mathbb{C}\setminus\mathbb{R}$ and $\hat{\tau}_1 \in \mathbb{R}$ be the zeroes of the *cubic* polynomial $(\tau-\hat{\alpha}_0/2)^2\prod_{j=1}^3(\tau-\hat{\alpha}_j)-(\tau-\hat{\beta}_0/2)^2\prod_{j=1}^3(\tau-\hat{\beta}_j)$. Define

$$\begin{aligned}\hat{f}_0(\tau) &= \hat{\alpha}_0\log(\tau-\hat{\alpha}_0/2)-\hat{\beta}_0\log(\tau-\hat{\beta}_0/2)+\textstyle\sum_{j=1}^3\big(\hat{\alpha}_j\log(\tau-\hat{\alpha}_j)-\hat{\beta}_j\log(\tau-\hat{\beta}_j)\big)\\ &\quad -(\hat{\alpha}_0-\hat{\beta}_0)\log(\hat{\alpha}_0/2-\hat{\beta}_0/2)-(\hat{\alpha}_1-\hat{\beta}_1)\log(\hat{\alpha}_1-\hat{\beta}_1)\\ &\quad +(\hat{\beta}_2-\hat{\alpha}_2)\log(\hat{\beta}_2-\hat{\alpha}_2)+(\hat{\beta}_3-\hat{\alpha}_3)\log(\hat{\beta}_3-\hat{\alpha}_3).\end{aligned}$$

Then
$$\limsup_{n\to\infty}\frac{\log|\hat{r}_n|}{n}=\operatorname{Re}\hat{f}_0(\hat{\tau}_0)\quad\text{and}\quad \lim_{n\to\infty}\frac{\log|\hat{q}_n|}{n}=\operatorname{Re}\hat{f}_0(\hat{\tau}_1).$$
Furthermore,
$$\hat{\Phi}_n^{-1}\hat{q}_n,\ 2\hat{\Phi}_n^{-1}D_{\hat{\gamma}_1 n}D_{\hat{\gamma}_2 n}^2\hat{p}_n\in\mathbb{Z}$$
with
$$\hat{\Phi}_n=\prod_{\substack{p\ \text{prime}\\ p\le\min\{\hat{\gamma}_1,\hat{\gamma}_2\}n}}p^{\hat{\varphi}(n/p)},$$
where
$$\begin{aligned}\hat{\varphi}(x)=\min_{0\le y<1}\Big(& \lfloor 2y-\hat{\beta}_0x\rfloor-\lfloor 2y-\hat{\alpha}_0x\rfloor-\lfloor(\hat{\alpha}_0-\hat{\beta}_0)x\rfloor\\ & +\lfloor y-\hat{\beta}_1x\rfloor-\lfloor y-\hat{\alpha}_1x\rfloor-\lfloor(\hat{\alpha}_1-\hat{\beta}_1)x\rfloor\\ & +\lfloor(\hat{\beta}_2-\hat{\alpha}_2)x\rfloor-\lfloor\hat{\beta}_2x-y\rfloor-\lfloor y-\hat{\alpha}_2x\rfloor\\ & +\lfloor(\hat{\beta}_3-\hat{\alpha}_3)x\rfloor-\lfloor\hat{\beta}_3x-y\rfloor-\lfloor y-\hat{\alpha}_3x\rfloor\Big)\end{aligned}$$
so that we have
$$\lim_{n\to\infty}\frac{\log\hat{\Phi}_n}{n}=\int_0^1\hat{\varphi}(x)\,\mathrm{d}\psi(x)-\int_0^{1/\min\{\hat{\gamma}_1,\hat{\gamma}_2\}}\hat{\varphi}(x)\,\frac{\mathrm{d}x}{x^2}.$$

11 Simultaneous diophantine properties of $\zeta(2)$ and $\zeta(3)$

In this section we prove Theorem 9.1 stated in section 9 (which exactly Theorem 3.1 stated in the introduction of this thesis page 17) by combining the constructions of Subsect. 10.2 and 10.3.

Construction 1. If we specialize the set of parameters (a, b) of Subsect. 10.2 to be
$$\begin{aligned}&a_1=8n+1, && a_2=7n+1, && a_3=10n+1, && a_4=9n+1,\\ &b_1=1, && b_2=n+1, && b_3=2n+1, && b_4=15n+2,\end{aligned}\tag{42}$$
then Propositions 10.5 and 10.6 imply that
$$r_n=r(a,b)=q_n\zeta(2)-p_n,\qquad\text{where}\quad \Phi_n^{-1}q_n,\ \Phi_n^{-1}D_{8n}D_{16n}p_n\in\mathbb{Z},\tag{43}$$

and

$$q_n = \frac{(-1)^n\,(9n)!\,(10n)!}{n!\,(2n)!\,(3n)!\,(5n)!\,(8n)!}\,{}_4F_3\left(\begin{matrix}-5n,\ 10n+1,\ 9n+1,\ 8n+1\\ 3n+1,\ 2n+1,\ n+1\end{matrix}\,\middle|\,1\right). \tag{44}$$

The corresponding function $\varphi(x)$ which defines Φ_n is

$$\varphi(x) = \begin{cases} 1 & \text{if } x \in \left[\frac{1}{10},\frac{1}{9}\right) \cup \left[\frac{1}{7},\frac{2}{9}\right) \cup \left[\frac{2}{7},\frac{1}{3}\right) \cup \left[\frac{2}{5},\frac{1}{2}\right) \cup \left[\frac{5}{9},\frac{4}{7}\right) \cup \left[\frac{2}{3},\frac{5}{7}\right) \cup \left[\frac{4}{5},\frac{6}{7}\right), \\ 2 & \text{if } x \in \left[\frac{1}{9},\frac{1}{8}\right) \cup \left[\frac{2}{9},\frac{1}{4}\right) \cup \left[\frac{1}{3},\frac{3}{8}\right) \cup \left[\frac{4}{7},\frac{5}{8}\right) \cup \left[\frac{5}{7},\frac{3}{4}\right) \cup \left[\frac{6}{7},\frac{7}{8}\right), \\ 0 & \text{otherwise}, \end{cases}$$

so that

$$\lim_{n\to\infty} \frac{\log \Phi_n}{n} = 6.61268356\ldots,$$

and the growth of r_n and q_n as $n \to \infty$ is determined by

$$\limsup_{n\to\infty} \frac{\log|r_n|}{n} = -19.10095491\ldots \quad \text{and} \quad \lim_{n\to\infty} \frac{\log|q_n|}{n} = 27.86755317\ldots.$$

Construction 2. If we specialize the set of parameters $(\hat{a},\hat{b})$ of Subsect. 10.3 to be

$$\begin{aligned} &\hat{a}_0 = 16n+2, \quad \hat{a}_1 = 8n+1, \quad \hat{a}_2 = 9n+1, \quad \hat{a}_3 = 10n+1, \\ &\hat{b}_0 = 11n+2, \quad \hat{b}_1 = 1, \qquad\quad \hat{b}_2 = 16n+2, \quad \hat{b}_3 = 16n+2, \end{aligned} \tag{45}$$

we obtain from Propositions 10.7 and 10.8 that

$$\hat{r}_n = \hat{r}(\hat{a},\hat{b}) = \hat{q}_n\zeta(3) - \hat{p}_n, \qquad \text{where} \quad \hat{\Phi}_n^{-1}\hat{q}_n,\ 2\hat{\Phi}_n^{-1}D_{8n}^3\hat{p}_n \in \mathbb{Z}, \tag{46}$$

and

$$\hat{q}_n = \frac{(7n)!\,(9n)!\,(10n)!}{n!\,(2n)!\,(4n)!\,(5n)!\,(6n)!\,(8n)!}\,{}_5F_4\left(\begin{matrix}-6n,\ -6n,\ 10n+1,\ \frac{9}{2}n+\frac{1}{2},\ \frac{9}{2}n+1\\ 2n+1,\ n+1,\ 2n+\frac{1}{2},\ 2n+1\end{matrix}\,\middle|\,1\right). \tag{47}$$

The corresponding function $\hat{\varphi}(x)$ assumes the form

$$\hat{\varphi}(x) = \begin{cases} 1 & \text{if } x \in \left[\frac{1}{10},\frac{1}{8}\right) \cup \left[\frac{1}{7},\frac{1}{4}\right) \cup \left[\frac{2}{7},\frac{1}{3}\right) \cup \left[\frac{3}{7},\frac{1}{2}\right) \cup \left[\frac{5}{9},\frac{4}{7}\right) \cup \left[\frac{3}{5},\frac{5}{8}\right) \cup \left[\frac{2}{3},\frac{5}{7}\right) \cup \left[\frac{5}{6},\frac{6}{7}\right), \\ 2 & \text{if } x \in \left[\frac{1}{3},\frac{3}{8}\right) \cup \left[\frac{4}{7},\frac{3}{5}\right) \cup \left[\frac{5}{7},\frac{3}{4}\right) \cup \left[\frac{6}{7},\frac{7}{8}\right), \\ 0 & \text{otherwise}, \end{cases}$$

so that

$$\lim_{n\to\infty} \frac{\log \hat{\Phi}_n}{n} = \varphi = 5.70169601\ldots,$$

and the growth of $\hat{r}_n$ and $\hat{q}_n$ as $n \to \infty$ is determined by

$$\limsup_{n\to\infty} \frac{\log|\hat{r}_n|}{n} = -\rho = -19.10095491\ldots \quad \text{and} \quad \lim_{n\to\infty} \frac{\log|\hat{q}_n|}{n} = \kappa = 27.86755317\ldots$$

with the same letters φ, κ and ρ as in section 9 and in the introduction of this thesis.

Connection between the constructions. Surprisingly — and this could be guessed from the asymptotics above, the coefficients in (43) of $\zeta(2)$ and in (46) of $\zeta(3)$ coincide: $q_n = \hat{q}_n$. This follows from the following classical identity — Whipple's transformation [Sla66, p. 65, eq. (2.4.2.3)], in which we assume that $b = -N$ is a negative integer:

$$
\begin{aligned}
{}_4F_3\left(\begin{matrix} f,\, 1+f-h,\, h-a,\, b \\ h,\, 1+f+a-h,\, g \end{matrix} \,\middle|\, 1\right) &= \frac{(g-f)_N}{(g)_N} \\
&\quad\times {}_5F_4\left(\begin{matrix} a, b, & 1+f-g, & \frac{1}{2}f, & \frac{1}{2}f+\frac{1}{2} \\ h, & 1+f+a-h, & \frac{1}{2}(1+f+b-g), & \frac{1}{2}(1+f+b-g)+\frac{1}{2} \end{matrix} \,\middle|\, 1\right). \qquad (48)
\end{aligned}
$$

The particular choices (42) and (45) correspond to taking $a = b = -6n$, $f = 9n+1$, $h = n+1$ and $g \to -n+1$ in (48). The equality $q_n = \hat{q}_n$ can be alternatively established by examining the recurrence equation satisfied by both q_n and $\hat{q}_n$; we outline the equation in our proof of Theorem 9.1 below.

Note that we also have Φ_n divisible by $\hat{\Phi}_n$ in the construction above, so that we can 'merge' the corresponding arithmetic properties (43) and (46) as follows:

$$\hat{\Phi}_n^{-1} q_n,\ \hat{\Phi}_n^{-1} D_{8n} D_{16n} p_n,\ 2\hat{\Phi}_n^{-1} D_{8n}^3 \hat{p}_n \in \mathbb{Z}. \qquad (49)$$

In both situations we get

$$\lim_{n\to\infty} \frac{\log(\hat{\Phi}_n^{-1} D_{8n} D_{16n})}{n} = \lim_{n\to\infty} \frac{\log(2\hat{\Phi}_n^{-1} D_{8n}^3)}{n} = 24 - \varphi = 18.29830398\ldots$$

and

$$\lim_{n\to\infty} \frac{\log|\hat{q}_n|}{n} = \kappa = 27.86755317\ldots,$$

so that both families of rational approximations to $\zeta(2)$ and $\zeta(3)$ are diophantine:

$$
\begin{aligned}
\limsup_{n\to\infty} \frac{\log|\hat{\Phi}_n^{-1} D_{8n} D_{16n} r_n|}{n} &= \limsup_{n\to\infty} \frac{\log|2\hat{\Phi}_n^{-1} D_{8n}^3 \hat{r}_n|}{n} \\
&= 24 - \varphi - \rho = -0.80265093\ldots < 0.
\end{aligned}
$$

Proof of Theorem 9.1. Using the notation above we define τ_0 and s_0 in accordance with (14).

To prove the theorem, we use a recurrence relation satisfied by q_n, p_n and $\hat{p}_n$. We execute the Gosper–Zeilberger algorithm of creative telescoping separately for the rational function $R_n(t) = R(t)$ defined in (18) and specialised by (42), and for $\hat{R}_n(t) = \hat{R}(t)$ defined in (30) with the choice of parameters (45). The results in both cases are polynomials $P_0(n), \ldots, P_3(n) \in \mathbb{Z}[n]$ and rational functions $S_n(t), \hat{S}_n(t)$ such that

$$
\begin{aligned}
P_3(n)R_{n+3}(t) + P_2(n)R_{n+2}(t) + P_1(n)R_{n+1}(t) + P_0(n)R_n(t) &= S_n(t+1) - S_n(t), \\
P_3(n)\hat{R}_{n+3}(t) + P_2(n)\hat{R}_{n+2}(t) + P_1(n)\hat{R}_{n+1}(t) + P_0(n)\hat{R}_n(t) &= \hat{S}_n(t+1) - \hat{S}_n(t).
\end{aligned}
$$

Applying then the argument as in the proof of Theorem 5.4 in [BBBC07] we find out that both the hypergeometric integrals

$$r_n = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\left(\frac{\pi}{\sin \pi t}\right)^2 R_n(t)\,\mathrm{d}t \quad \text{and} \quad \hat{r}_n = \frac{1}{4\pi i}\int_{-i\infty}^{i\infty}\left(\frac{\pi}{\sin \pi t}\right)^2 \hat{R}_n(t)\,\mathrm{d}t$$

satisfy the *same* recurrence equation

$$P_3(n)y_{n+3} + P_2(n)y_{n+2} + P_1(n)y_{n+1} + P_0(n)y_n = 0.$$

Since $r_n = q_n\zeta(2) - p_n$, $\hat{r}_n = q_n\zeta(3) - \hat{p}_n$ and both $\zeta(2)$ and $\zeta(3)$ are irrational, we deduce that the coefficients q_n, p_n and $\hat{p}_n$ satisfy the same equation. Using this fact we obtain that the sequence of determinants

$$\Delta_n = \begin{vmatrix} q_n & q_{n+1} & q_{n+2} \\ p_n & p_{n+1} & p_{n+2} \\ \hat{p}_n & \hat{p}_{n+1} & \hat{p}_{n+2} \end{vmatrix}$$

satisfies the recurrence equation $P_3(n)\Delta_{n+1} + P_0(n)\Delta_n = 0$. The coefficients of $P_3(n)$ are all positive, while the coefficients of $P_0(n)$ are all negative. On the webpage of the first author, one can find the polynomials P_0, P_1, P_2 and P_3 and the proof of the claim on the sign of their coefficients, the computation of the linear forms r_n and $\hat{r}_n$ for the first values of n, the computation of the determinant below and the computation of the recurrence relation verified by r_n and $\hat{r}_n$ together with the proof of their equality. More details on how these computations have been done can be found in section 13 below.

The relation $P_3(n)\Delta_{n+1} + P_0(n)\Delta_n = 0$, together with the fact that $P_0(n)$ and $P_3(n)$ do not vanish, implies that the nonvanishing of Δ_n for some n is equivalent to the nonvanishing of Δ_0. We have explicitly

$$\begin{aligned}
&q_0 = 1, \quad && q_1 = 12307565655, \quad && q_2 = 566993126516654178841 5, \\
&p_0 = 0, \quad && p_1 = \tfrac{199536684432021}{9856}, \quad && p_2 = \tfrac{6500408024275547867356589727409007}{696970391040}, \\
&\hat{p}_0 = 0, \quad && \hat{p}_1 = \tfrac{7953492001094261}{537600}, \quad && \hat{p}_2 = \tfrac{37762843816152998347068580008855083}{5540664729600},
\end{aligned}$$

so that

$$\Delta_0 = \begin{vmatrix} q_0 & q_1 & q_2 \\ p_0 & p_1 & p_2 \\ \hat{p}_0 & \hat{p}_1 & \hat{p}_2 \end{vmatrix} = \tfrac{28866666573725618155283921483481 9523}{1072688684225235517 44000} \neq 0\,.$$

Thus, $\Delta_n \neq 0$ for any $n \geq 0$.

Now let ε, $\eta > 0$; for simplicity we may assume $\eta \leq \varepsilon$. Let m be a sufficiently large integer as in the statement of Theorem 9.1. Let a_0, a_1, a_2 satisfy the hypotheses in Theorem 9.1. We take $n = \lceil m/8 \rceil$, so that $8n - 7 \leq m \leq 8n$. Since the determinant Δ_n does not vanish, there exists an $\ell \in \{n\,, n+1\,, n+2\}$ such that

$$a_0 q_\ell + a_1 p_\ell + a_2 \hat{p}_\ell \neq 0\,.$$

Now we have $m \leq 8n \leq 8\ell$, so that $D_{8\ell}^2 D_{16\ell} a_0 \in \mathbb{Z}$ and $D_{8\ell} a_1 \in \mathbb{Z}$. Letting $e_{m,\ell} = \frac{2D_{8\ell}}{D_m}$, we get the property

$$\frac{D_{2m}}{D_m} \,\Big|\, e_{m,l} \frac{D_{16\ell}}{2D_{8\ell}},$$

so that $e_{m,\ell} \dfrac{D_{16\ell}}{2D_{8\ell}} a_2 \in \mathbb{Z}$. Therefore, using the arithmetic properties of q_ℓ, p_ℓ and $\hat{p}_\ell$ we conclude that

$$e_{m,\ell}(D_{8\ell}^2 D_{16\ell} a_0)(\Phi_\ell^{-1} q_\ell) + e_{m,\ell}(D_{8\ell} a_1)(\Phi_\ell^{-1} D_{8\ell} D_{16\ell} p_\ell) + e_{m,\ell}\left(\frac{D_{16\ell}}{2D_{8\ell}} a_2\right)(2\Phi_\ell^{-1} D_{8\ell}^3 \hat{p}_\ell) \quad (50)$$

is a nonzero integer. Note that $\ell \leq \frac{m}{8} + 3$, so that the asymptotic contribution of $e_{m,\ell}$ is almost invisible: $e_{m,\ell} \leq \frac{2D_{m+24}}{D_m} = e^{o(m)} = e^{o(\ell)}$.

Let us bound the integer (50) from above. Writing hypothesis (ii) as

$$|a_0 + a_1\zeta(2) + a_2\zeta(3)| \leq e^{-(s_0+\eta)m} \leq e^{-(32-\varphi+\kappa+8\eta)(n-1)},$$

we obtain

$$\begin{aligned} &|a_0 q_\ell + a_1 p_\ell + a_2 \hat{p}_\ell| \\ &\quad \leq |q_\ell|\,|a_0 + a_1\zeta(2) + a_2\zeta(3)| + |a_1|\,|q_\ell\zeta(2) - p_\ell| + |a_2|\,|q_\ell\zeta(3) - \hat{p}_\ell| \\ &\quad \leq e^{-(32-\varphi+8\varepsilon)n+o(n)}, \end{aligned}$$

since $\varepsilon \leq \eta$. On the other hand, the common denominator of the coefficients used above is

$$e_{m,\ell} D_{8\ell}^2 D_{16\ell} \Phi_\ell^{-1} \leq e^{(2\cdot 8+16-\varphi)\ell+o(\ell)} = e^{(32-\varphi)n+o(n)}.$$

This means that the non-zero integer (50) has absolute value at most $e^{-8\varepsilon n+o(n)}$, which is not possible for a sufficiently large n, thus implying the truth of Theorem 9.1. □

A second proof could have been done by applying a slightly more general version of Theorem 6.1 of section 6 page 31 of this Ph.D. thesis. Indeed, the form of the divisors $\delta_{2,n} = \frac{D_{16n}}{2D_{8n}}$ does not give us the divisiblity $\delta_{2,n} \big| \delta_{2,n+1}$. Therefore, we need to counterbalance this lack of divisibility by changing assumption (i) of Theorem 6.1 to $\delta_{i,n} \big| \delta_{i,n+1} e_n$ for some not too large e_n (as in the proof of Theorem 9.1). And then, the theorem could be applied.

Similarly, a slightly more general version of Theorem 5.1 of § 5 page 26 could also be applied here. With this theorem, the assumption (ii) on the asymptotic behaviors of the linear forms is not satisfied any more, since the linear forms oscillate. Using Fischler's work [Fis12] on Sorokin's trick ([Sor07], see also Remark 2 in §2.2 of [Fis12]) in addition to the same assumption $\delta_{i,n} | \delta_{i,n+1} e_n$ as above, one could derive Theorem 9.1.

12 A new diophantine exponent

12.1 Definition and basic properties

We now introduce a new exponent that depends on some $\tau \in \mathbb{R}$ and is related to Theorem 9.1.

Definition 12.1 :
Let $\xi_1, \xi_2 \in \mathbb{R}$ and $\tau \in \mathbb{R}$. We denote by $s_\tau(\xi_1, \xi_2)$ the infimum of the set $E_\tau(\xi_1, \xi_2)$ of all $s \in \mathbb{R}$ with the following property. Let $\varepsilon > 0$ and n be sufficiently large in terms of ε. Let $(a_0, a_1, a_2) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ be such that:

(i) $D_n^2 D_{2n} a_0 \in \mathbb{Z}$, $D_n a_1 \in \mathbb{Z}$ and $\dfrac{D_{2n}}{D_n} a_2 \in \mathbb{Z}$; and

(ii) $|a_0|, |a_1|, |a_2|$ are bounded from above by $e^{-(\tau+\varepsilon)n}$.

Then $|a_0 + a_1\xi_1 + a_2\xi_2| > e^{-sn}$.
By convention, we set $s_\tau(\xi_1, \xi_2) = +\infty$ if $E_\tau(\xi_1, \xi_2) = \emptyset$, and $s_\tau(\xi_1, \xi_2) = -\infty$ if $E_\tau(\xi_1, \xi_2) = \mathbb{R}$.

This definition allows us to restate Theorem 9.1 as follows.

Theorem 12.1 :
With $\tau_0 = 0.899668635\ldots$ and $s_0 = 6.770732145\ldots$ as in (14), we have $s_{\tau_0}(\zeta(2), \zeta(3)) \le s_0$.

To begin with, let us state and prove general results on this diophantine exponent $s_\tau(\xi_1, \xi_2)$ depending on the range in which τ varies; it turns out that it carries diophantine information on ξ_1 and ξ_2 only if $\tau < 1$.

Proposition 12.2 :

1. If $\tau > 4$, then $s_\tau(\xi_1, \xi_2) = -\infty$.
2. If $1 \le \tau \le 4$, then $s_\tau(\xi_1, \xi_2) = 4$.
3. If $\tau < 1$, then $s_\tau(\xi_1, \xi_2) \ge 6 - 2\tau$.
4. If $\tau < 1$ and at least one of ξ_1 or ξ_2 is rational, then $s_\tau(\xi_1, \xi_2) = +\infty$.
5. If $\tau < 0$ and the numbers 1, ξ_1 and ξ_2 are linearly dependent over $\mathbb{Q}$, then $s_\tau(\xi_1, \xi_2) = +\infty$.
6. If $\tau \le \tau'$, then $s_\tau(\xi_1, \xi_2) \ge s_{\tau'}(\xi_1, \xi_2)$.

Proof. (1) We see that whenever the coefficient a_i is not zero, we must have $|a_i| \ge 1/(D_n^2 D_{2n}) = e^{-4n+o(n)}$ if $i = 0$, and an even larger estimate from below (namely, $e^{-n+o(n)}$) if $i = 1$ or 2. Therefore, having at least one triple $(a_0, a_1, a_2) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ that satisfies both (i) and (ii) of Definition 12.1 yields $\tau \le 4$; having no such triple implies $E_\tau(\xi_1, \xi_2) = \mathbb{R}$.

(2) Assuming now $1 \le \tau \le 4$ in Definition 12.1 and choose n sufficiently large to accommodate $D_n < e^{(1+\varepsilon)n}$ and $D_{2n}/D_n < e^{(1+\varepsilon)n}$. Condition (ii) implies that $|a_1| \le e^{-(\tau+\varepsilon)n} \le e^{-n-\varepsilon n}$, so that the integer $|D_n a_1| \le D_n e^{-n-\varepsilon n} < 1$ must be zero, $a_1 = 0$. Similar consideration shows that $a_2 = 0$, hence the only nonzero element in the triple $(a_0, a_1, a_2) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ is a_0. Then condition (i) implies that $|a_0| \ge 1/(D_n^2 D_{2n}) = e^{-4n+o(n)}$ with the equality possible

by simply taking $a_0 = 1/(D_n^2 D_{2n})$. Thus, $s_\tau(\xi_1, \xi_2) = 4$ for all ξ_1, ξ_2 whenever $4 \ge \tau \ge 1$.

(3) Take $s < 6 - 2\tau$ and define $\varepsilon = \frac{1}{3}(6 - 2\tau - s)$, so that $s = 6 - 2\tau - 3\varepsilon > \tau + \varepsilon/2$ because of $\tau < 1$. Let n be sufficiently large to have $(D_n D_{2n})^2 > e^{(6-\varepsilon)n} = e^{-(s+2\tau+2\varepsilon)n}$ satisfied. Define the set

$$K = \{(x_0, x_1, x_2) \in \mathbb{R}^3 : |x_1|, |x_2| \le e^{-(\tau+\varepsilon)n},\ |x_0 + x_1\xi_1 + x_2\xi_2| \le e^{-sn}\} \subset \mathbb{R}^3,$$

which is compact, convex, symmetric with respect to $\mathbf{0}$ and has volume $8e^{-(s+2\tau+2\varepsilon)n}$. Consider the lattice

$$\Gamma = \frac{1}{D_n^2 D_{2n}}\mathbb{Z} \oplus \frac{1}{D_n}\mathbb{Z} \oplus \frac{D_n}{D_{2n}}\mathbb{Z},$$

whose fundamental domain has volume

$$\frac{1}{D_n^2 D_{2n}} \cdot \frac{1}{D_n} \cdot \frac{D_n}{D_{2n}} < e^{-(s+2\tau+2\varepsilon)n}.$$

By Minkowski's theorem, K contains a nonzero point (a_0, a_1, a_2) of the lattice Γ, for which we have

$$\begin{aligned} |a_0| &\le |a_1|\,|\xi_1| + |a_2|\,|\xi_2| + |a_0 + a_1\xi_1 + a_2\xi_2| \\ &\le (|\xi_1| + |\xi_2|)e^{-(\tau+\varepsilon)n} + e^{-sn} \le e^{-(\tau+\varepsilon/2)n}. \end{aligned}$$

The estimate means that $s \notin E_\tau(\xi_1, \xi_2)$; as $s_\tau(\xi_1, \xi_2)$ is the infimum of the set $E_\tau(\xi_1, \xi_2)$, we get $s_\tau(\xi_1, \xi_2) \ge 6 - 2\tau$.

(4) Assume $\xi_1 = p/q \in \mathbb{Q}$, take $\varepsilon \in (0, 1 - \tau)$. By choosing $a_0 = q\xi_1/D_n$, $a_1 = -q/D_n$ and $a_2 = 0$ we see that properties (i) and (ii) in the definition of $E_\tau(\xi_1, \xi_2)$ are satisfied for *any* n sufficiently large. In addition, $|a_0 + a_1\xi_1 + a_2\xi_2| = 0 < e^{-sn}$ for any $s \in \mathbb{R}$, which means that $E_\tau(\xi_1, \xi_2) = \emptyset$, hence $s_\tau(\xi_1, \xi_2) = +\infty$.

If $\xi_2 = p/q \in \mathbb{Q}$, then the choice $a_0 = q\xi_2 D_n/D_{2n}$, $a_1 = 0$ and $a_2 = -qD_n/D_{2n}$ does the job.

(5) Assume now that there exist integers q_0, q_1 and q_2, not all zero, such that $p_0 + p_1\xi_1 + p_2\xi_2 = 0$. Setting $a_0 = p_0$, $a_1 = p_1$ and $a_2 = p_2$ we see that properties (i) and (ii) are satisfied with any choice of $\tau < 0$ and ε, for all n sufficiently large in terms of ε. At the same time $|a_0 + a_1\xi_1 + a_2\xi_2| = 0 < e^{-sn}$ for any $s \in \mathbb{R}$, meaning that $E_\tau(\xi_1, \xi_2) = \emptyset$, hence $s_\tau(\xi_1, \xi_2) = +\infty$.

(6) Using (1), (2) and (3), we may assume that $\tau' < 1$. Let $s \in E_\tau(\xi_1, \xi_2)$ meaning that for all $\varepsilon > 0$ and for all triples $(a_0, a_1, a_2) \in \mathbb{Q}^3 \backslash \{\mathbf{0}\}$ which satisfy $D_n^2 D_{2n} a_0$, $D_n a_1$, $\dfrac{D_{2n}}{D_n} a_2 \in \mathbb{Z}$ and $|a_0|, |a_1|, |a_2| \le e^{-(\tau+\varepsilon)n}$, we have $|a_0 + a_1\xi_1 + a_2\xi_2| > e^{-sn}$. For n be sufficiently large and $(a_0, a_1, a_2) \in \mathbb{Q}^3 \setminus \{\mathbf{0}\}$ such that $D_n^2 D_{2n} a_0$, $D_n a_1$, $\dfrac{D_{2n}}{D_n} a_2 \in \mathbb{Z}$ and $|a_i| \le e^{-(\tau'+\varepsilon)n}$, we also have $|a_i| \le e^{-(\tau+\varepsilon)n}$. This means that $|a_0 + a_1\xi_1 + a_2\xi_2| > e^{-sn}$ and $s \in E_{\tau'}(\xi_1, \xi_2)$, so that $E_\tau(\xi_1, \xi_2) \subset E_{\tau'}(\xi_1, \xi_2)$, which leads to claim (6) by taking the infimum of both sets. □

From now on we assume τ to be real < 1.

Remarks. Theorem 12.1 is nontrivial since $\tau_0 < 1$. However, it does not imply that 1, $\zeta(2)$ and $\zeta(3)$ are $\mathbb{Q}$-linearly independent since $\tau_0 > 0$.

Part (3) of Proposition 12.2 yields $s_{\tau_0}(\zeta(2), \zeta(3)) \geq 4.20$, so that the statement of Theorem 12.1 is far from being best possible.

The fact that $s_{\tau_0}(\zeta(2), \zeta(3)) < +\infty$ in Theorem 12.1 is already new.

12.2 Omitting one number

Recall the definition of the usual exponent of irrationality of $\mu(\xi)$ of a number $\xi \in \mathbb{R}$ from the introductory part. Here comes its generalisation, the *ψ-exponent of irrationality*, given by S. Fischler in [Fis09].

Definition 12.2 :
Let $\mathcal{E}$ be the set of all $\psi\colon \mathbb{N}^* \to \mathbb{N}^*$ with the following properties: for any $q \geq 1$, $\psi(q+1)$ is a multiple of $\psi(q)$, and the limit

$$\gamma_\psi = \lim_{q\to\infty} \frac{\log \psi(q)}{\log q}$$

exists and belongs to the interval $[0, 1)$. For $\psi \in \mathcal{E}$ and $\xi \in \mathbb{R} \setminus \mathbb{Q}$, denote by $\mu_\psi(\xi)$ the supremum of the set $M_\psi(\xi)$ of all $\mu \in \mathbb{R}$ such that there are infinitely many $q \geq 1$ which are divisible by $\psi(q)$ and satisfy

$$\left|\xi - \frac{p}{q}\right| \leq \frac{1}{q^\mu} \quad \text{for some } p \in \mathbb{Z}\,.$$

If $M_\psi(\xi)$ is not bounded from above, that is, if $M_\psi(\xi) = \mathbb{R}$, we get $\mu_\psi(\xi) = +\infty$.

An equivalent way of defining $\mu_\psi(\xi)$, is by letting $\mu_\psi(\xi)$ be the infimum of the set of exponents μ such that for all q large enough with $\psi(q) \mid q$ one has $|\xi - p/q| > 1/q^{-\mu}$, and taking $\mu_\psi(\xi) = +\infty$ if the set is empty.

When $\psi(q) = 1$ for all q, the ψ-exponent $\mu_\psi(\xi)$ coincides with the usual exponent of irrationality $\mu(\xi)$. It is known [Fis09, Corollary 3] that $\mu_\psi(\xi) = +\infty$ if and only if ξ is a Liouville number, that is, $\mu(\xi) = +\infty$. If this is not the case, then

$$(1 - \gamma_\psi)\mu(\xi) \leq \mu_\psi(\xi) \leq \mu(\xi).$$

Fischler proves in [Fis09] that $\mu_\psi(\xi) \geq 2 - \gamma_\psi$ for any $\psi \in \mathcal{E}$ and any $\xi \in \mathbb{R} \setminus \mathbb{Q}$, with the equality holding for almost all $\xi \in \mathbb{R}$ in the sense of Lebesgue measure. More precisely, he shows that, given an $\eta > 2 - \gamma_\psi$, the set of ξ such that $\mu_\psi(\xi) > \eta$ has Hausdorff dimension $(2 - \gamma_\psi)/\eta$.

The usual construction of a function $\psi \in \mathcal{E}$ is as follows. One takes $\psi(q) = \delta_n$ with $n = \lfloor (\log q)/(\delta - \alpha) \rfloor$, where $(\delta_n)_{n\geq 1}$ is a sequence of positive integers such that δ_n divides δ_{n+1} for each $n \geq 1$ and $\delta_n = e^{\delta n + o(n)}$ as $n \to \infty$, while $\alpha \in \mathbb{R}$ is chosen to satisfy $\alpha < \delta$. In this construction, we have $\gamma_\psi = \delta/(\delta - \alpha)$.

Definition 12.2 allows us to deduce diophantine results involving only one quantity, ξ_1 or ξ_2, from a nontrivial upper bound on the exponent $s_\tau(\xi_1, \xi_2)$ from Definition 12.1.

Proposition 12.3 :
Let ξ_1, ξ_2 be real numbers and $\tau < 1$. Define $\psi_1, \psi_2 \colon \mathbb{N}^* \to \mathbb{N}^*$ by taking $\psi_1(q) = D_n D_{2n}$ and $\psi_2(q) = D_n^3$, where $n = \lfloor (\log q)/(4-\tau) \rfloor$. Then

$$\mu_{\psi_i}(\xi_i) \le \frac{s_\tau(\xi_1, \xi_2) - \tau}{4 - \tau} \quad \text{for } i = 1, 2.$$

Proof. Let $\tau' \in \mathbb{R}$ satisfy $\tau < \tau' < 1$. Take $p \in \mathbb{Z}$ and $q \in \mathbb{N}^*$ sufficiently large, $\psi_1(q) \mid q$, and $m = \lfloor (\log q/(4-\tau') \rfloor$, so that $\psi_1(q) = D_m D_{2m}$. We may assume that $|p/q - \xi_1| < 1$.

For an $s > s_\tau(\xi_1, \xi_2)$, choose $\varepsilon > 0$ such that

$$\varepsilon < \frac{1}{2} \min\left\{ \tau' - \tau \, , \, \frac{(s-4)(\tau' - \tau)}{2s - \tau' - 4} \right\};$$

part (3) of Proposition 12.2 implies $s > 4$.

Take

$$n = \left\lfloor \frac{4 - \tau'}{4 - \tau - 2\varepsilon}(m+1) \right\rfloor + 1 < m, \quad a_0 = \frac{p}{D_n^2 D_{2n}} \quad \text{and} \quad a_1 = \frac{-q}{D_n^2 D_{2n}}.$$

Then

$$D_n^2 D_{2n} a_0 \in \mathbb{Z} \quad \text{and} \quad D_n a_1 = \frac{-q}{D_m D_{2m}} \frac{D_m D_{2m}}{D_n D_{2n}} = \frac{-q}{\psi_1(q)} \frac{D_m D_{2m}}{D_n D_{2n}} \in \mathbb{Z},$$

and $q < e^{(4-\tau')(m+1)}$ implying that $|a_1| \le e^{-(\tau+\varepsilon)n}$; for a_0 we have $|a_0| = |p| e^{-4n + o(n)}$. Therefore, $|p| \le |p - q\xi_1| + |q\xi_1| \le q(1 + |\xi_1|)$, which leads to

$$|a_0| \le q(1 + |\xi_1|) e^{-4n + o(n)} \le e^{(4-\tau')(m+1) - (4-\varepsilon)n} \le e^{-(\tau+\varepsilon)n}$$

for q sufficiently large. Letting $a_2 = 0$ and using $s > s_\tau(\xi_1, \xi_2)$ we deduce that $|a_0 + a_1\xi_1| > e^{-sn}$; from the definition of a_0 and a_1 it follows that

$$\left| \frac{p}{q} - \xi_1 \right| > \frac{e^{(4-s-\varepsilon)n}}{q}$$

provided q is sufficiently large. The assumption on ε results in the estimate

$$\left| \frac{p}{q} - \xi_1 \right| > q^{-(s-\tau')/(4-\tau')},$$

which implies $\mu_{\psi_1}(\xi_1) \le (s - \tau')/(4 - \tau')$. This upper bound holds for all $s > s_\tau(\xi_1, \xi_2)$; taking the infimum over s and then choosing $\tau' \in (\tau, 1)$ sufficiently close to τ, completes the proof for $i = 1$. The proof for $i = 2$ is similar. $\square$

Since $\mu_{\psi_i}(\xi_i) \geq 2 - \gamma_{\psi_i}$ with $\gamma_{\psi_i} = 3/(4-\tau) < 1$, Proposition 12.3 implies the lower bound $s_\tau(\xi_1, \xi_2) \geq 5 - \tau$, which is however weaker than the one from statement (3) of Proposition 12.2.

Corollary 12.4 :
For ξ_1 and ξ_2 real numbers and $\tau < 1$, the following inequalities hold for the ordinary irrationality exponent:

$$\mu(\xi_i) \leq \frac{s_\tau(\xi_1, \xi_2) - \tau}{1 - \tau} \quad \text{for } i = 1, 2.$$

Proof. In the notation of Proposition 12.3, use $(1 - \gamma_{\psi_i})\mu(\xi_i) \leq \mu_{\psi_i}(\xi_i)$. □

If we use the upper bound of Theorem 12.1 together with Proposition 12.4 above, one can deduce the upper bounds $\mu(\zeta(2)), \mu(\zeta(3)) \leq 58.516731\ldots$ which are much worse than the upper bounds obtained by Rhin-Viola recalled in the introduction of this chapter.

12.3 Case of linear dependence

In this section, we prove a converse result to Proposition 12.3 above, namely we deduce (if $1, \xi_1$ and ξ_2 are linearly dependent over $\mathbb{Q}$) an upper bound on $s_\tau(\xi_1, \xi_2)$ from an upper bound on an irrationality exponent which depends only on ξ_1 (resp. ξ_2).

Lemma 12.5 :
Let $\alpha \geq 1$ and ψ be a function as in Definition 12.2.
Then, for all $\varepsilon > 0$, and any integer n large enough in terms of ε such that $\psi(n) \big| n$, there exists $k \geq 1$ such that $k \leq n^\varepsilon$ and $\psi(k\alpha n) \big| k\alpha n$.

Proof. Let $k = \frac{\psi([\alpha n^{1+\varepsilon}])}{\psi(n)}$ which is an integer since $n \leq \alpha n^{1+\varepsilon}$. From the definition of γ_ψ, we know that $\psi(n) = n^{\gamma_\psi + o(1)}$. And as $[\alpha n^{1+\varepsilon}] = \alpha n^{1+\varepsilon} + O(1)$, we also have $\psi([\alpha n^{1+\varepsilon}]) = (\alpha n^{1+\varepsilon} + O(1))^{\gamma_\psi + o(1)} = n^{(1+\varepsilon)\gamma_\psi + o(1)}$. Thus,

$$k = \frac{n^{(1+\varepsilon)\gamma_\psi + o(1)}}{n^{\gamma_\psi + o(1)}} = n^{\varepsilon\gamma_\psi + o(1)}\,.$$

Since $0 \leq \gamma_\psi < 1$, if n is large enough, we deduce that $k < n^\varepsilon$. Therefore, $k\alpha n < \alpha n^{1+\varepsilon}$, and finally $\psi(k\alpha n) \big| \psi([\alpha n^{1+\varepsilon}]) = k\psi(n) \big| kn \big| k\alpha n$. □

Lemma 12.6 :
If $1, \xi_1, \xi_2$ are $\mathbb{Q}$-linearly dependent with $\xi_1, \xi_2 \notin \mathbb{Q}$ and ψ is a function as in Definition 12.2, then

$$\mu_\psi(\xi_1) = \mu_\psi(\xi_2)\,.$$

Proof. Let $(\alpha_0, \alpha_1, \alpha_2) \in \mathbb{Z}^3 \setminus \{0\}$ be such that $\alpha_0 + \alpha_1\xi_1 + \alpha_2\xi_2 = 0$. As ξ_1 and ξ_2 are not rationals, we know that both α_1 and α_2 are non-zero integers.

By symmetry, it is sufficient to prove that $\mu_\psi(\xi_2) \leq \mu_\psi(\xi_1)$, and we may assume $\alpha_1 < 0$.

Since $\xi_1 \notin \mathbb{Q}$, we have $\mu_\psi(\xi_1) \geq 2 - \gamma_\psi > 1$; let $\varepsilon > 0$ be such that $\varepsilon < \mu_\psi(\xi_1) - 1$. Let $p \in \mathbb{Z}$ and $q \in \mathbb{N}^*$ be such that $\psi(q)\big|q$ and q is large enough. Then

$$\left|\frac{p}{q} - \xi_2\right| = \left|\frac{p}{q} + \frac{\alpha_0}{\alpha_2} + \frac{\alpha_1}{\alpha_2}\xi_1\right| = \left|\frac{\alpha_1}{\alpha_2}\right|\left|\frac{p\alpha_2 + q\alpha_0}{\alpha_1 q} + \xi_1\right| = \left|\frac{\alpha_1'}{\alpha_2'}\right|\left|\frac{p\alpha_2' + q\alpha_0'}{\alpha_1' q} - \xi_1\right|$$

with $\alpha_i' = -\alpha_i$ for $i = 0, 1, 2$.

Then, we take k to be the integer given by Lemma 12.5 applied to $\alpha_1' \geq 0$ and $n = q$. Thus $\left|\frac{p}{q} - \xi_2\right| = \left|\frac{\alpha_1'}{\alpha_2'}\right|\left|\frac{k(p\alpha_2' + q\alpha_0')}{k\alpha_1' q} - \xi_2\right|$, with $\psi(k\alpha_1' q)\big|k\alpha_1' q$.

Letting $\mu = \mu_\psi(\xi_1)$, we deduce, since $k\alpha_1' q \geq q$ is sufficiently large with respect to ε, that:

$$\left|\frac{p}{q} - \xi_2\right| > \left|\frac{\alpha_1'}{\alpha_2'}\right| \frac{1}{(k\alpha_1' q)^{\mu+\varepsilon}} > \frac{1}{q^{\mu(1+2\varepsilon)}}$$

since $k < q^\varepsilon$, q is large enough in terms of ε, and α_1 and α_2 are fixed non-zero integers depending only on ξ_1 and ξ_2 and $(\mu + \varepsilon)(1 + \varepsilon) < \mu(1 + 2\varepsilon)$ because $1 + \varepsilon < \mu$.

We have proved that $\mu_\psi(\xi_2) \leq (1 + 2\varepsilon)\mu_\psi(\xi_1)$; taking ε arbitrarily small yields $\mu_\psi(\xi_2) \leq \mu_\psi(\xi_1)$, which concludes the proof of the lemma (as noticed above). $\square$

Proposition 12.7 :
Let $\xi_1, \xi_2 \notin \mathbb{Q}$ and assume that $1, \xi_1, \xi_2$ are $\mathbb{Q}$-linearly dependent. Let $0 \leq \tau < 1$ and let $\widetilde{\psi}$ be the map from $\mathbb{N}^*$ to $\mathbb{N}^*$ defined by $\widetilde{\psi}(q) = D_n^2$ with $n = \left\lfloor \frac{\log q}{4-\tau} \right\rfloor$.
Then,

$$s_\tau(\xi_1, \xi_2) \leq 4 + (\mu_{\widetilde{\psi}}(\xi_i) - 1)(4 - \tau)$$

for $i = 1, 2$.
In addition,

$$s_\tau(\xi_1, \xi_2) \leq 6 - \tau_0$$

unless both ξ_1 and ξ_2 belong to a certain set of Lebesgue measure zero.

We remark that the conclusion does not hold if $\tau < 0$ since $s_\tau(\xi_1, \xi_2) = +\infty$ in this case (see Proposition 12.2).

Proof. Let $\alpha_0, \alpha_1 \in \mathbb{Q}$ be such that $\alpha_0 + \alpha_1\xi_1 = \xi_2$. Let $\varepsilon > 0$, $\nu > 0$, $\mu > \mu_{\widetilde{\psi}}(\xi_1)$ and n be sufficiently large with respect to ε, ν and μ. Let $(a_0, a_1, a_2) \in \mathbb{Q}^3$ with $(a_0, a_1, a_2) \neq 0$, $|a_i| \leq e^{-(\tau+\varepsilon)n}$ and the arithmetic properties $D_n^2 D_{2n} a_0$, $D_n a_1$ and $\frac{D_{2n}}{D_n} a_2 \in \mathbb{Z}$. Let $\eta = |a_0 + a_1\xi_1 + a_2\xi_2|$. We denote by A a common denominator of α_0 and α_1.

To begin with, we claim that $\eta \neq 0$. Indeed, if $\eta = 0$ then we have $a_0 = -\alpha_0 a_2$ and $a_1 = -\alpha_1 a_2$, since 1, ξ_1, ξ_2 span a $\mathbb{Q}$-vector space of dimension 2 (so that there is exactly one $\mathbb{Q}$-linear relation among them, up to proportionality). The rational number $a_1 = -\alpha_1 a_2$ admits both D_n and $A\frac{D_{2n}}{D_n}$ as denominator, so that $\delta_n a_1 \in \mathbb{Z}$, where $\delta_n = \gcd(D_n, A\frac{D_{2n}}{D_n})$. Assume that $a_1 \neq 0$; then $\delta_n |a_1|$ is a positive integer so that $\delta_n \geq |a_1|^{-1} \geq e^{(\tau+\varepsilon)n} \geq e^{\varepsilon n}$, since $\tau \geq 0$. This is a contradiction because $\gcd(D_n, \frac{D_{2n}}{D_n}) = e^{o(n)}$ as $n \to \infty$. Therefore, $a_1 = 0$, so that $a_2 = 0$ since $\alpha_1 \neq 0$ (because $\xi_2 \notin \mathbb{Q}$), and finally $0 = \eta = |a_0|$, which is impossible since $(a_0, a_1, a_2) \neq (0, 0, 0)$. This concludes the proof of the claim that $\eta \neq 0$.

We write now $\eta = |a'_0 + a'_1\xi_1|$ with $a'_0 = a_0 + \alpha_0 a_2$ and $a'_1 = a_1 + \alpha_1 a_2$.

If $a'_1 \neq 0$, we have the arithmetic properties $AD_n^2 D_{2n} a'_0 \in \mathbb{Z}$ and $AD_{2n} a'_1 \in \mathbb{Z}$. Set $a''_0 = -\mathrm{sign}(a'_1) AD_n^2 D_{2n} a'_0 \in \mathbb{Z}$ and $a''_1 = AD_n^2 D_{2n} |a'_1| \in D_n^2\mathbb{N}$. By assumption, we know that $a''_1 > 0$ which leads us to the inequalities $e^{2n+o(n)} \leq a''_1 \leq e^{(4-\tau-\varepsilon)n+o(n)} \leq e^{(4-\tau)n}$. Thus, $\frac{\log a''_1}{4-\tau} \leq n$ which ensures us to have $\widetilde{\psi}(a''_1) \big| D_n^2 | a''_1$. Since $a''_1 \geq e^{2n+o(n)}$ and n is large enough in terms of μ with $\mu > \mu_{\widetilde{\psi}}(\xi_1)$, we obtain:

$$AD_n^2 D_{2n}\eta = |a''_0 - a''_1\xi_1| > \frac{1}{a''^{\mu-1}_1} > e^{-(\mu-1)(4-\tau)n}$$

which yields $\eta > e^{-\left(4+(\mu-1)(4-\tau)+\nu\right)n}$ if n is large enough, since A is a fixed integer depending only on ξ_1 and ξ_2.

Now, let us consider the case $a'_1 = 0$, so that $\eta = |a'_0|$. Since $\eta \neq 0$, $AD_n^2 D_{2n} a_0 \in \mathbb{Z}$ and $A\frac{D_{2n}}{D_n} a_2 \in \mathbb{Z}$, this implies $AD_n^2 D_{2n}\eta \in \mathbb{N}^*$ and thus $\eta > e^{-4n+o(n)}$. But $\gamma_{\widetilde{\psi}} \in [0, 1)$ leads to $\mu_{\widetilde{\psi}}(\xi_1) > 1$ with the help of the inequality $\mu_{\widetilde{\psi}}(\xi_1) \geq 2 - \gamma_{\widetilde{\psi}}$. Thus, we know that $(\mu_{\widetilde{\psi}}(\xi_1) - 1)(4-\tau) > 0$. Hence, if n is large enough, we have $\eta > e^{-\left(4+(\mu-1)(4-\tau)+\nu\right)n}$ in this case too.

Therefore, in both cases, we have $4 + (\mu-1)(4-\tau) + \nu \in E_\tau(\xi_1, \xi_2)$ for all $\mu > \mu_{\widetilde{\psi}}(\xi_1)$ and all $\nu > 0$. But as $s_\tau(\xi_1, \xi_2)$ is the infimum of the set $E_\tau(\xi_1, \xi_2)$, the desired inequality follows with $i = 1$. It holds also with $i = 2$ using Lemma 12.6.

Finally $\mu_\psi(\xi) = 2 - \gamma_\psi = 2 - \frac{2}{4-\tau}$ for almost all $\xi \in \mathbb{R}$ with respect ot the Lebesgue measure, completes the proof. □

12.4 Rational approximation to $\zeta(3)$ only

Combining Theorem 12.1 with Proposition 12.3, we deduce the following result.

Proposition 12.8 :
For $\psi(q) = D_n^3$ with $n = \lfloor (\log q)/(4-\tau_0) \rfloor$ and τ_0 defined in (14), we have the upper bound

$$\mu_\psi(\zeta(3)) \leq 1.92357696\ldots.$$

Let us conclude with a few remarks on this result.

As shown in [Fis09], Apéry's proof of the irrationality of $\zeta(3)$ leads to the estimate $\mu_{\psi'}(\zeta(3)) \leq 2$, where $\psi'(q) = D_n^3$ with $n = \lfloor (\log q)/(4\log(1+\sqrt{2})) \rfloor$. Since $4\log(1+\sqrt{(2)}) > 4 - \tau_0$, this implies $\mu_\psi(\zeta(3)) \leq 2$ with the function ψ in Proposition 12.8. Therefore, Proposition 12.8 is slightly sharper than what follows from Apéry's construction.

Proposition 12.8 can be adapted to $\zeta(2)$; namely, we have $\mu_{\tilde{\psi}}(\zeta(2)) \leq 1.92$, where $\tilde{\psi}(q) = D_n D_{2n}$ with $n = \lfloor (\log q)/(4-\tau_0) \rfloor$ and τ_0 as before. This result is slightly sharper than the one that can be deduced from Apéry's construction. Apéry's proof yields $\mu_{\tilde{\psi}'}(\zeta(2)) \leq 2$, where $\tilde{\psi}'(q) = D_n^2$ with $n = \lfloor (\log q)/(5(\log(1+\sqrt{5}) - \log 2)) \rfloor$. Using Lemma 12.9 below, this upper bound can be shown to imply $\mu_{\tilde{\psi}}(\zeta(2)) \leq 3.103$, which is greater than the one from Proposition 12.8.

In the notation above, and again with the help of Lemmas 12.9 and 12.10 below, the upper bound of Proposition 12.8 and its analogue for $\zeta(2)$ imply $\mu_{\tilde{\psi}'}(\zeta(2)) \leq 15.54$ and $\mu_{\psi'}(\zeta(3)) \leq 8.85$: these upper bounds are worse than the ones followed from Apéry's construction.

Proposition 12.8 means that $\zeta(3)$ does not belong to the set of $\xi \in \mathbb{R} \setminus \mathbb{Q}$ satisfying $\mu_\psi(\xi) > 1.92\ldots$. This set has Hausdorff dimension equal to $0.0681457\ldots$; this is smaller than the one obtained after Corollary 5 in [Fis09].

Finally, for the function $\psi \in \mathcal{E}$ in Proposition 12.8 we have that $\mu_\psi(\xi) = 1.03\ldots$ for almost all $\xi \in \mathbb{R}$. Therefore, Proposition 12.8 is still quite far from being optimal, since $\zeta(3)$ is presumably a 'generic' real number.

Lemma 12.9 :
Let $\xi \in \mathbb{R} \setminus \mathbb{Q}$. Let ψ_1 and ψ_1' be defined by $\psi_1(q) = D_n D_{2n}$ and $\psi_1'(q) = D_{n'}^2$ with $n = \left\lfloor \frac{\log q}{4-\tau_0} \right\rfloor$ and $n' = \left\lfloor \frac{\log q}{5(\log(1+\sqrt{5}) - \log 2)} \right\rfloor$. Let

$$K_1 = \frac{5(1-\tau_0)(\log(1+\sqrt{5}) - \log 2)}{(3-\tau_0)(5\log(1+\sqrt{5}) - \log 2) - 4 + \tau_0}$$

and

$$K_2 = \frac{5(4-\tau_0)(\log(1+\sqrt{5}) - \log 2) - 5(\log(1+\sqrt{5}) - \log 2) - 4 + \tau_0}{5(4-\tau_0)(\log(1+\sqrt{5}) - \log 2) - 2(4-\tau_0)}.$$

Then we have the inequalities:

$$K_1 \mu_{\psi_1'}(\xi) \leq \mu_{\psi_1}(\xi) \leq K_2 \mu_{\psi_1'}(\xi)$$

with $K_1 = 0.1235\ldots$ and $K_2 = 1.55148\ldots$.

Lemma 12.10 :
Let $\xi \in \mathbb{R} \setminus \mathbb{Q}$. Let ψ_2 and ψ_2' be defined by $\psi_2(q) = 2D_n^3$ and $\psi_2'(q) = D_{n'}^3$ with $n = \left\lfloor \frac{\log q}{4-\tau_0} \right\rfloor$ and $n' = \left\lfloor \frac{\log q}{4\log(1+\sqrt{2})} \right\rfloor$. Then we have the inequalities:

$$\frac{4(1-\tau_0)\log(1+\sqrt{2})}{(4-\tau_0)(4\log(1+\sqrt{2}) - 3)} \mu_{\psi_2'}(\xi) \leq \mu_{\psi_2}(\xi) \leq \mu_{\psi_2'}(\xi),$$

with

$$\frac{4(1-\tau_0)\log(1+\sqrt{2})}{(4-\tau_0)(4\log(1+\sqrt{2}) - 3)} = 0.21711\ldots.$$

Proof of Lemmas 12.9 and 12.10. Let us begin with the proof of Lemma 12.10. Since $\frac{1}{4\log(1+\sqrt{2})} < \frac{1}{4-\tau_0}$, we have

$$n' = \left\lfloor \frac{\log q}{4\log(1+\sqrt{2})} \right\rfloor \leq n = \left\lfloor \frac{\log q}{4-\tau_0} \right\rfloor$$

for q large enough, so that $\psi_2'(q) \big| \psi_2(q)$. Therefore, Lemma 2 of [Fis09] yields

$$\frac{1-\gamma_{\psi_2}}{1-\gamma_{\psi_2'}} \mu_{\psi_2'}(\xi) \leq \mu_{\psi_2}(\xi) \leq \mu_{\psi_2'}(\xi)\,,$$

with $\gamma_{\psi_2'} = \frac{3}{4\log(1+\sqrt{2})}$ and $\gamma_{\psi_2} = \frac{3}{4-\tau_0}$. This concludes the proof of Lemma 12.10.

Let us prove Lemma 12.9. From $\frac{4-\tau_0}{5(\log(1+\sqrt{5})-\log 2)} = 1.28855\ldots \in (1,2)$, we get

$$n = \frac{\log q}{4-\tau_0} + O(1) \leq n' = \frac{\log q}{5(\log(1+\sqrt{5})-\log 2)} + O(1) \leq 2n = \frac{2\log q}{4-\tau_0} + O(1)$$

with $n = \left\lfloor \frac{\log q}{4-\tau_0} \right\rfloor$ and $n' = \left\lfloor \frac{\log q}{5(\log(1+\sqrt{5})-\log 2)} \right\rfloor$. Thus $n \leq n' \leq 2n$ and so $D_n \big| D_{n'} \big| D_{2n}$ which implie that $D_n D_{n'}$ divides the gcd of $D_n D_{2n}$ and $D_{n'}^3$. Now, define ψ_3 by $\psi_3(q) = D_n D_{n'}$ with these values of n and n'. Thus, $\psi_3(q) = e^{n+n'+o(n')}$. Hence we have $\gamma_{\psi_3} = \frac{1}{4-\tau_0} + \frac{1}{5(\log(1+\sqrt{5})-\log 2)} = 0.73816\ldots$. So ψ_3 verifies all the requirements of Definition 12.2. We can apply Lemma 2 of [Fis09] again to ψ_3 and ψ_1 and then to ψ_3 and ψ_1' in order to have respectively the comparisons:

$$\frac{1-\gamma_{\psi_1}}{1-\gamma_{\psi_3}} \mu_{\psi_3}(\xi) \leq \mu_{\psi_1}(\xi) \leq \mu_{\psi_3}(\xi)$$

and

$$\frac{1-\gamma_{\psi_1'}}{1-\gamma_{\psi_3}} \mu_{\psi_3}(\xi) \leq \mu_{\psi_1'}(\xi) \leq \mu_{\psi_3}(\xi)\,.$$

Thus:

$$\frac{1-\gamma_{\psi_1}}{1-\gamma_{\psi_3}} \mu_{\psi_1'}(\xi) \leq \frac{1-\gamma_{\psi_1}}{1-\gamma_{\psi_3}} \mu_{\psi_3}(\xi) \leq \mu_{\psi_1}(\xi) \leq \mu_{\psi_3}(\xi) \leq \frac{1-\gamma_{\psi_3}}{1-\gamma_{\psi_1'}} \mu_{\psi_1'}(\xi)$$

which leads to Lemma 12.9.

□

13 Computations in the proof of Theorem 9.1

In this section, details about the computations made for the proof of Theorem 9.1 in § 11 are given. Here, the computations discussed are the one used in the direct proof of Theorem 9.1, the one that does not involve independence criteria of sections 5 and 6. All the notation of part III are used here.

First, some recurrence relations verified by the linear forms r_n and $\hat{r}_n$ and their coefficients are needed. To find them, we applied the creative telescoping algorithm of Zeilberger. The user instructions for this algorithm are fully explained in the book [PWZ96]. A version of this algorithm is freely available on the book's webpage. This algorithm was implemented for various softwares, like Mathematica or Maple. In the present case, Maple was used for convenience and personal interest and also because Zeilberger's algorithm is already included in Maple (at least since version 9). All the computations described here are available in both Maple format .mw and texte format .txt on the webpage `http://www.math.u-psud.fr/~dauguet/These/These.html`.

The creative telescoping algorithm applies to an hypergeometric expression $F(n,k)$ in n and k, meaning an expression such that $F(n+1,k)/F(n,k)$ and $F(n,k+1)/F(n,k)$ are both rational functions in n and k. After loading the hypergeometric package with `with(SumTools[Hypergeometric])`, it is called by `Zeilberger(F(n,k),n,k,N)`. The input `N` is the name of the shift operator with respect to the varaible n, namely the operator such that, for a sequence $u = (u_n)$, $(Nu)_n = u_{n+1}$. The output is a list of 2 elements $[L, G]$, where L is polynomial in n and N and G a function of the variables n and k. They are linked by the telescoping relation $L(n,N)F(n,k) = G(n,k+1) - G(n,k)$. The function G is actually a certificate to check the well computation of the algorithm. It allows one, in principle, to prove the recurrence relation without the help of a computer by checking this identity directly.

But in the present computations, we used the command `IsZapplicable`, since the command `Zeilberger` doesn't seem to work on our version of Maple. The command `IsZapplicable` has either 3 or 5 entries. The first 3 are necessary and must be `F`,`n` and `k`, which are the hypergeometric term F depending on n and k. The command determines if the creative telescoping algorithm of Zeilberger is applicable or not to F. It returns either `True` or `False`. The last two entries are optional and are respectively `N`, the name of the shift operator, and `'Name'`, a string (remark that the quote marks `'` are necessary here). If the 5 inputs are given and in the case where Zeilberger's algorithm is indeed appliable to F, the computation of the algorithm is done and the result of this computation is put in the variable `Name`. The output of `IsZapplicable(F,n,k,N,'Name')` is still `True` or `False`. In shorter words, the command `IsZapplicable(F,n,k,N,'Name')` is equivalent to the command `IsZapplicable(F,n,k)` followed by `Name:=Zeilberger(F,n,k,N)` if the answer is `True`, so that `Name` contains both the recurrence relation and the certificate as a list.

The rational fractions $R(t)$ and $\hat{R}(t)$ of § 10 with the choice of parameters of § 11 are hypergeometric functions in n and t. Therefore Zeilberger's creative telescoping algorithm can be applied and it gave the polynomials $P_0, \ldots, P_3$ used in the proof, which are the coefficients of $N^0, \ldots, N^3$ in $L(n,N)$ (and $L(n,N)$ has degree 3 in N). These polynomials $P_0(n), \ldots, P_3(n)$ are fully displayed on the webpage dedicated to this Ph.D. Thesis (`http://www.math.u-psud.`

fr/~dauguet/These/These.html). Since they have degree 100 and coefficients about 10^{113}, we have chosen to trust the algorithm and didn't take much interest in the certificate G.

The recurrence relation given by $P_0, \ldots, P_3$ is also verified by r_n and $\hat{r}_n$ since they are expressed linearly in terms of the rational fractions $R_n(t)$ and $\hat{R}_n(t)$. The hypergeometric forms of the coefficients q_n and $\hat{q}_n$ of r_n and $\hat{r}_n$ given in (44) and (47) in § 11 allow us also to apply Zeilberger's algorithm to it. We get also the same recurrence relation, so that all coefficients p_n, $\hat{p}_n$, q_n, $\hat{q}_n$, the linear forms r_n and $\hat{r}_n$ verify the same recurrence relation and so that one can deduce the equality $q_n = \hat{q}_n$ claimed in § 11 from the recurrence relation.

And finally, we needed to compute the determinant

$$\begin{vmatrix} q_0 & q_1 & q_2 \\ p_0 & p_1 & p_2 \\ \hat{p}_0 & \hat{p}_1 & \hat{p}_2 \end{vmatrix}$$

to check whether it is 0 or not. The point here was of course the exact computation of the values of p_n and $\hat{p}_n$, since q_n is readily computed using (44) or (47) (see page 63), since the two are equal.

We use the relations (20) page 54 and (23) page 55 to have the values of the remainder $P(t)$ on the integers in the set $[-8n, 8n-1]$. These values allow us to decompose P in the basis of polynomials P_ℓ according to the first formulae of the proof of Proposition 10.5 page 56, by inverting a matrix. Then, we did some verification of our computation. Finally, we compute the coefficients p_n from the formulae of r_n given in (27) page 56.

We did then the same procedure for $\hat{p}_n$. From the explicit formula of $\hat{r}_n$ given at the beginning of the proof of Proposition 10.7 in formula (40) page 60, the coefficients $\hat{p}_n$ are completely determined by the coefficients A_k and B_k. The explicit formula (33) page 59 gives us an easy way to compute the A_i's. And the formula (34) page 59 for B_k is simple enough to be computed easily. We did also some verifications of the computations, for safety.

And last but not least, we finally were able to compute the determinant above.

Bibliographie

[Apé79] Roger APÉRY : Irrationalité de $\zeta(2)$ et $\zeta(3)$. *Astérisque*, 61:11–13, 1979.

[BBBC07] David H. BAILEY, David BORWEIN, Jonathan M. BORWEIN et Richard E. CRANDALL : Hypergeometric forms for Ising-class integrals. *Experiment. Math.*, 16(3): 257–276, 2007.

[Bed98] Egor V. BEDULEV : On the linear independence of numbers over number fields. *Mat. Zametki*, 64(4):506–517, 1998.

[Bel10] Pierre BEL : Fonctions L p-adiques et irrationalité. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 9(1):189–227, 2010.

[BR01] Keith BALL et Tanguy RIVOAL : Irrationalité d'une infinité de valeurs de la fonction zêta aux entiers impairs. *Invent. Math.*, 146(1):193–207, 2001.

[Bun08] Peter BUNDSCHUH : Arithmetical results on certain q-series. I. *Int. J. Number Theory*, 4(1):25–43, 2008.

[Bun12] Peter BUNDSCHUH : Transcendence and algebraic independence of series related to Stern's sequence. *Int. J. Number Theory*, 8(2):361–376, 2012.

[BV11] Peter BUNDSCHUH et Keijo VÄÄNÄNEN : An application of Nesterenko's dimension estimate to p-adic q-series. *Int. J. Number Theory*, 7(2):431–447, 2011.

[Cas97] John W. S. CASSELS : *An introduction to the geometry of numbers*. Classics in Mathematics. Springer-Verlag, Berlin, 1997.

[Cha12] Amarisa CHANTANASIRI : Généralisation des critères pour l'indépendance linéaire de Nesterenko, Amoroso, Colmez, Fischler et Zudilin. *Ann. Math. Blaise Pascal*, 19(1):75–105, 2012.

[CZ08] Stephen CHOI et Ping ZHOU : On linear independence of a certain multivariate infinite product. *Canad. Math. Bull.*, 51(1):32–46, 2008.

[Dau14] Simon DAUGUET : Généralisations quantitatives du critère d'indépendance linéaire de Nesterenko. HAL :hal-00959764, arXiv : 1403.4205 . Soumis, 2014.

[DZ14] Simon Dauguet et Wadim Zudilin : On simultaneous diophantine approximations to $\zeta(2)$ and $\zeta(3)$. arXiv :1401.5322, HAL :hal-00933967. Soumis, 2014.

[FHKL13] Stéphane Fischler, Mumtaz Hussain, Simon Kristensen et Jason Levesley : A converse to linear independence criteria, valid almost everywhere. arXiv :1302.1952, Soumis, 2013.

[Fis09] Stéphane Fischler : Restricted rational approximation and Apéry-type constructions. *Indagationes Mathematicae*, 20 (2):201–215, Juin 2009.

[Fis12] Stéphane Fischler : Nesterenko's criterion when the small linear forms oscillate. *Arch. Math. (Basel)*, 98(2):143–151, 2012.

[Fis13] Stéphane Fischler : Nesterenko's linear independence criterion for vectors. arXiv :1202.2279. Soumis, 2013.

[FN98] Naum I. Fel'dman et Yuri V. Nesterenko : Transcendental numbers. *In Number theory, IV*, volume 44 de *Encyclopaedia Math. Sci.*, pages 1–345. Springer, Berlin, 1998.

[FR10] Stéphane Fischler et Tanguy Rivoal : Irrationality exponent and rational approximations with prescribed growth. *Proc. Amer. Math. Soc.*, 138(3):799–808, 2010.

[FZ10] Stéphane Fischler et Wadim Zudilin : A refinement of Nesterenko's linear independence criterion with applications to zeta values. *Math. Annalen*, 347:739–763, 2010.

[Hat95] Masayoshi Hata : A note on Beukers' integral. *J. Austral. Math. Soc. Ser. A*, 58(2):143–153, 1995.

[Hat00] Masayoshi Hata : A new irrationality measure for $\zeta(3)$. *Acta Arith.*, 92(1):47–57, 2000.

[HKO12] Noriko Hirata-Kohno et Hironori Okada : A note on linear independence of polylogarithms over the rationals. *Proc. Japan Acad. Ser. A Math. Sci.*, 88(9):156–161, 2012.

[Lin82] Ferdinand Lindemann : Ueber die Zahl π. *Math. Ann.*, 20(2):213–225, 1882.

[Mah39] Kurt Mahler : Ein Übertragungsprinzip für konvexe Körper. *Časopis Pěst. Mat. Fys.*, 68:93–102, 1939.

[Mar06] Raffaele Marcovecchio : Linear independence of linear forms in polylogarithms. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)*, 5(1):1–11, 2006.

[Nes85] Yuri V. Nesterenko : Linear independence of numbers. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, (1):46–49, 108, 1985.

[Nes12] Yuri V. NESTERENKO : On a criterion of linear independence of p-adic numbers. *Manuscripta Math.*, 139(3-4):405–414, 2012.

[Phi86] Patrice PHILIPPON : Critères pour l'indépendance algébrique. *Inst. Hautes Études Sci. Publ. Math.*, (64):5–52, 1986.

[PS98] George PÓLYA et Gabor SZEGŐ : *Problems and theorems in analysis. II.* Classics in Mathematics. Springer-Verlag, Berlin, 1998.

[PWZ96] Marko PETKOVŠEK, Herbert S. WILF et Doron ZEILBERGER : $A = B$. A K Peters Ltd., Wellesley, MA, 1996.

[Rie59] Bernhard RIEMANN : Ueber die Anzahl des Primzahlen unter einer gegebenen Grösse. *Monatsberichte des Berliner Akademie*, Novembre 1859. Traduction en anglais sur le site `http://www.maths.tcd.ie/pub/HistMath/People/Riemann/Zeta/`.

[Riv00] Tanguy RIVOAL : La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs. *C.R.A.S. Paris Série I Math.*, 336.4:267–270, 2000.

[Riv02] Tanguy RIVOAL : Irrationalité d'au moins un des neuf nombres $\zeta(5),\zeta(7),\ldots,\zeta(21)$. *Acta Arith.*, 103(2):157–167, 2002.

[Riv12] Tanguy RIVOAL : On the arithmetic nature of the values of the gamma function, Euler's constant, and Gompertz's constant. *Michigan Math. J.*, 61(2):239–254, 2012.

[RV96] Georges RHIN et Carlo VIOLA : On a permutation group related to $\zeta(2)$. *Acta Arith.*, 77(1):23–56, 1996.

[RV01] Georges RHIN et Carlo VIOLA : The group structure for $\zeta(3)$. *Acta Arith.*, 97(3): 269–293, 2001.

[Sie29] Carl L. SIEGEL : Über einige Anwendungen Diophantischer Approximationen. *Abh. Press. Akad. Wiss. Phys.-Math. Kl.*, 1:1–70, 1929. JFM 56.0180.01.

[Sla66] Lucy J. SLATER : *Generalized hypergeometric functions*. Cambridge University Press, Cambridge, 1966.

[Sor07] Vladimir N. SOROKIN : On the Zudilin-Rivoal theorem. *Mat. Zametki*, 81(6):912–923, 2007.

[Töp94] Thomas TÖPFER : An axiomatization of Nesterenko's method and applications on Mahler functions. *J. Number Theory*, 49(1):1–26, 1994.

[Töp95] Thomas TÖPFER : An axiomatization of Nesterenko's method and applications on Mahler functions. II. *Compositio Math.*, 95(3):323–342, 1995.

[Vas97] Oleg N. VASILENKO : On the approximation of differences of dilogarithms and their values at rational points. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, (1):10–12, 78, 1997.

[Vä10] Keijo O. VÄÄNÄNEN : Remarks on the linear independence of some q-series. *Fundam. Prikl. Mat.*, 16(5):41–47, 2010.

[Zud01] Wadim ZUDILIN : One of the numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational. *Russian Math. Surveys*, 56.4:774–776, 2001.

[Zud02] Wadim ZUDILIN : Irrationality of values of the Riemann zeta function. *Izv. Math.*, 3:489–542, 2002.

[Zud04] Wadim ZUDILIN : Arithmetic of linear forms involving odd zeta values. *J. Théorie des Nombres de Bordeaux*, 16(1):251–291, 2004.

[Zud07] Wadim ZUDILIN : Approximations to -, di- and tri- logarithms. *J. Comput. Appl. Math.*, 202(2):450–459, 2007.

[Zud11] Wadim ZUDILIN : Arithmetic hypergeometric series. *Uspekhi Mat. Nauk*, 66(2(398)):163–216, 2011.

[Zud13] Wadim ZUDILIN : On the irrationality measure of π^2. *Uspekhi Mat. Nauk*, 68(6): 171–172 (in Russian), 2013. Taduction en anglais dans Russian Mat. Surveys 68(6), 2013, à paraître.

[Zud14] Wadim ZUDILIN : Two hypergeometric tales and a new irrationality measure of $\zeta(2)$. *Ann. Math. Québec*, 38, 2014. arXiv :1310.1526.